

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 13:08:05 ; Search time 123 Seconds
(without alignments)
6501.498 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfilesi.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	2.5	2255	3	US-08-871-572B-3
2	35	2.4	1639	3	US-09-362-473-5
3	34	2.4	144	1	US-08-702-344-26
4	34	2.4	348	4	US-09-621-976-13740
5	34	2.4	388	4	US-09-621-976-18573
6	34	2.4	396	4	US-09-640-173-10
7	34	2.4	396	4	US-09-713-550-10
8	34	2.4	397	4	US-09-621-976-18571
9	34	2.4	769	4	US-08-567-882-5
10	34	2.4	1191	3	US-09-282-305-13
11	34	2.4	1191	4	US-09-883-720-13
12	34	2.4	1283	1	US-08-174-467-19
13	34	2.4	1283	3	US-08-452-071-19
14	34	2.4	1518	4	US-09-614-912-191
15	34	2.4	1582	3	US-08-545-196B-10
16	34	2.4	1582	3	US-08-545-196B-12
17	34	2.4	2043	4	US-09-614-912-171
18	34	2.4	2103	4	US-09-489-847-40
19	34	2.4	2218	2	US-08-985-090-4
20	34	2.4	2218	3	US-09-165-543-31
21	34	2.4	2780	4	US-09-489-847-87
22	34	2.4	3244	3	US-09-165-543-4
23	34	2.4	3437	3	US-08-860-339-17
24	34	2.4	3437	4	US-09-573-629-17
25	34	2.4	3581	2	US-08-738-349-1
26	34	2.4	4494	4	US-09-620-312D-861
27	33	2.3	229	4	US-09-621-976-10914

28	33	2.3	273	4	US-09-621-976-11078	Sequence 11078, A
29	33	2.3	375	3	US-08-946-026-23	Sequence 23, Appl
30	33	2.3	724	3	US-09-020-956-38	Sequence 38, Appl
31	33	2.3	724	3	US-09-030-607-38	Sequence 38, Appl
32	33	2.3	724	4	US-09-439-313-38	Sequence 38, Appl
33	33	2.3	724	4	US-09-352-616A-38	Sequence 38, Appl
34	33	2.3	724	4	US-09-232-149A-38	Sequence 38, Appl
35	33	2.3	724	4	US-09-159-812-38	Sequence 38, Appl
36	33	2.3	724	4	US-09-636-215-38	Sequence 38, Appl
37	33	2.3	724	4	US-09-685-166A-38	Sequence 38, Appl
38	33	2.3	724	4	US-09-115-453-38	Sequence 38, Appl
39	33	2.3	724	4	US-08-924-759-9	Sequence 9, Appl
40	33	2.3	911	2	US-08-248-335-9	Sequence 9, Appl
41	33	2.3	1052	4	US-09-489-847-23	Sequence 23, Appl
42	33	2.3	1223	3	US-09-154-874-4	Sequence 4, Appl
43	33	2.3	1223	4	US-08-931-558-4	Sequence 4, Appl
44	33	2.3	1223	4	US-09-468-175-4	Sequence 4, Appl
45	33	2.3	1223	4	US-09-468-175-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-871-572B-3
; Sequence 3, Application US/08871572B
; Patent No. 6287853
; GENERAL INFORMATION:
; APPLICANT: Pestka, Sidney
; APPLICANT: Kotenko, Serguei
; APPLICANT: Soh, Jaemog
; APPLICANT: Donnelly, Robert
; APPLICANT: Mariano, Thomas
; APPLICANT: Cook, Jeffrey
; APPLICANT: Emmanuel, Stuart
; APPLICANT: Schwartz, Barbara
; TITLE OF INVENTION: Accessory Factor for Interferon Gamma
; TITLE OF INVENTION: and Its Receptor
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: 758 Springfield Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,572B
; FILING DATE: 9-JUNE-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Muccino, Richard R.
; REGISTRATION NUMBER: 32,538
; REFERENCE/DOCKET NUMBER: UMD1-011
; TELEPHONE: (908) 273-4988
; TELEFAX: (908) 273-4679
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-871-572B-3

Query Match 2.5%; Score 36; DB 3; Length 2255;
 Best Local Similarity 100.0%; Pred. No. 2.5e-06;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1406 TGTTTTTAAAAA 1441
 Db 2209 TGTTTTTAAAAA 2244

RESULT 2

US-09-362-473-5
 ; Sequence 5, Application US/09362473
 ; Patent No. 6218169
 ; GENERAL INFORMATION:
 ; APPLICANT: Cakoon, Edgar B.
 ; APPLICANT: Cakoon, Rebecca E.
 ; APPLICANT: Falco, S. Carl
 ; APPLICANT: Morgante, Michele
 ; APPLICANT: Rafalski, J. Antoni
 ; APPLICANT: Hitz, William D.
 ; APPLICANT: Kinney, Anthony J.
 ; TITLE OF INVENTION: Aromatic Amino Acid Catabolism Enzymes
 ; FILE REFERENCE: BB-1197
 ; CURRENT APPLICATION NUMBER: US/09/362,473
 ; CURRENT FILING DATE: 1999-07-28
 ; EARLIER APPLICATION NUMBER: 60/094,783
 ; EARLIER FILING DATE: JULY 31, 1998
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 5
 ; LENGTH: 1639
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; US-09-362-473-5

Query Match 2.4%; Score 35; DB 3; Length 1639;
 Best Local Similarity 100.0%; Pred. No. 7e-06;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1407 GTTTTTTAAAAA 1441
 Db 1591 GTTTTTTAAAAA 1625

RESULT 3

US-08-702-344-26
 ; Sequence 26, Application US/08702344
 ; Patent No. 5723315
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John
 ; APPLICANT: LaVallie, Edward
 ; APPLICANT: Racie, Lisa
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Spaulding, Vikki
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 ; TITLE OF INVENTION: ENCODING THEM
 ; NUMBER OF SEQUENCES: 37
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/702,344

; FILING DATE:
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brown, Scott A.
 ; REGISTRATION NUMBER: 32,724
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8224
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 144 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-702-344-26

Query Match 2.4%; Score 34; DB 1; Length 144;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTTTAAAAA 1441
 Db 17 TTTTTTAAAAA 50

RESULT 4

US-09-621-976-13740
 ; Sequence 13740, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 13740
 ; LENGTH: 348
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-621-976-13740

Query Match 2.4%; Score 34; DB 4; Length 348;
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTTTAAAAA 1441
 Db 314 TTTTTTAAAAA 347

RESULT 5

US-09-621-976-18573
 ; Sequence 18573, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Jobert, S.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
 ; FILE REFERENCE: GENSET.054PR2
 ; CURRENT APPLICATION NUMBER: US/09/621,976
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 19335
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 18573
 ; LENGTH: 388
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

US-09-621-976-18573

Query Match 2.4%; Score 34; DB 4; Length 388;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1441
|||||
Db 343 TTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 376
|||||

RESULT 6

US-09-640-173-10/c

; Sequence 10, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; FILE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640,173
; CURRENT FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G

US-09-640-173-10

Query Match 2.4%; Score 34; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1441
|||||
Db 110 TTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 77
|||||

RESULT 7

US-09-713-550-10/c

; Sequence 10, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713,550
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G

US-09-713-550-10

Query Match 2.4%; Score 34; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1441
|||||

Db 110 TTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 77
|||||

RESULT 8

US-09-621-976-18571

; Sequence 18571, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18571
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-621-976-18571

Query Match 2.4%; Score 34; DB 4; Length 397;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1441
|||||
Db 343 TTTTAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 376
|||||

RESULT 9

US-08-567-882-5

; Sequence 5, Application US/08567882
; Patent No. 6512103
; GENERAL INFORMATION:
; APPLICANT: Dairaghi, Daniel J.
; APPLICANT: Hara, Takahiko
; APPLICANT: Miyajima, Atsushi
; APPLICANT: Schall, Thomas J.
; APPLICANT: Wang, Wei
; APPLICANT: Yoshimura, Akihiko
; TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,882
; FILING DATE: 08-DEC-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0506
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-852-9196
; TELEFAX: 415-496-1200
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 769 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 US-08-567-882-5

Query Match 2.4%; Score 34; DB 4; Length 1191;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA 1441
 DB 1158 TTTTAAAAA 1191

RESULT 12
 US-08-174-467-19
 ; Sequence 19, Application US/08174467
 ; Patent No. 5451514
 ; GENERAL INFORMATION:
 ; APPLICANT: BOUDET, ALAIN M.
 ; APPLICANT: INZE, DIRK G.
 ; APPLICANT: SCHUCH, WOLFGANG W.
 ; TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN
 ; TITLE OF INVENTION: PLANTS
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
 ; STREET: 1100 NEW YORK AVENUE, N.W.
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/174,467
 ; FILING DATE: 28-DEC-1993
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/874,166
 ; FILING DATE: 27-APR-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KOKULIS, PAUL N.
 ; REGISTRATION NUMBER: 16,773
 ; REFERENCE/DOCKET NUMBER: 95563/PS36321/US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-861-3000
 ; TELEFAX: 202-822-0944
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1283 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA
 ; US-08-174-467-19

Query Match 2.4%; Score 34; DB 1; Length 1283;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA 1441
 DB 1244 TTTTAAAAA 1277

RESULT 13
 US-08-452-071-19
 ; Sequence 19, Application US/08452071
 ; Patent No. 6066780
 ; GENERAL INFORMATION:
 ; APPLICANT: BOUDET, ALAIN M.
 ; APPLICANT: INZE, DIRK G.
 ; APPLICANT: SCHUCH, WOLFGANG W.
 ; TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 US-08-567-882-5

Query Match 2.4%; Score 34; DB 4; Length 769;
 Best Local Similarity 100.0%; Pred. No. 2e-05;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA 1441
 DB 523 TTTTAAAAA 556

RESULT 10
 US-09-282-305-13
 ; Sequence 13, Application US/09282305
 ; Patent No. 6287843
 ; GENERAL INFORMATION:
 ; APPLICANT: Baldwin, Donald A.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Crane, Virginia C.
 ; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 ; FILE REFERENCE: 5718-44,
 ; CURRENT APPLICATION NUMBER: US/09/282,305
 ; CURRENT FILING DATE: 1999-03-31
 ; PRIOR APPLICATION NUMBER: 60/080,563
 ; PRIOR FILING DATE: 1998-04-03
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 13
 ; LENGTH: 1191
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (63)..(971)
 ; US-09-282-305-13

Query Match 2.4%; Score 34; DB 3; Length 1191;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA 1441
 DB 1158 TTTTAAAAA 1191

RESULT 11
 US-09-883-720-13
 ; Sequence 13, Application US/09883720
 ; Patent No. 6479629
 ; GENERAL INFORMATION:
 ; APPLICANT: Baldwin, Donald A.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Crane, Virginia C.
 ; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
 ; FILE REFERENCE: 5718-44,
 ; CURRENT APPLICATION NUMBER: US/09/883,720
 ; CURRENT FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: 09/282,305
 ; PRIOR FILING DATE: 1999-03-31
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 13
 ; LENGTH: 1191
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (63)..(971)
 ; US-09-883-720-13

TITLE OF INVENTION: PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,071
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,166
FILING DATE: 27-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 95563/PS36321/US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1283 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-452-071-19
Query Match 2.4%; Score 34; DB 3; Length 1283;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1408 TTTTAAAAA 1441
DB 1244 TTTTAAAAA 1277
RESULT 14
US-09-614-912-191
Sequence 191, Application US/09614912
Patent No. 6677502
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalski, Antoni
APPLICANT: Orozco, Buddy
APPLICANT: Miao, Gou-Hau
APPLICANT: Famodu, Omolayo O.
APPLICANT: Lee, Jian Ming
APPLICANT: Sakai, Hajime
APPLICANT: Weng, Zude
APPLICANT: Caimi, Perry G
APPLICANT: Anderson, Shawn
TITLE OF INVENTION: Plant Metabolism Genes
FILE REFERENCE: BB1378 US NA
CURRENT APPLICATION NUMBER: US/09/614,912
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 60/143,401
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/143,412
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/146,650
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/170,906
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: 60/172,959
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/172,946
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 204
SOFTWARE: Microsoft Office 97
SEQ ID NO 191
LENGTH: 1518
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (114)
NAME/KEY: unsure
LOCATION: (123)
NAME/KEY: unsure
LOCATION: (138)
US-09-614-912-191
Query Match 2.4%; Score 34; DB 4; Length 1518;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1408 TTTTAAAAA 1441
DB 1434 TTTTAAAAA 1467
RESULT 15
US-08-545-196B-10
Sequence 10, Application US/08545196B
Patent No. 6080577
GENERAL INFORMATION:
APPLICANT: MELKI, JUDITH
APPLICANT: MUNNICH, ARNOLD
TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,196B
FILING DATE: 19-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FARACI, C. J.
REGISTRATION NUMBER: 32,350
REFERENCE/DOCKET NUMBER: 2121-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-545-196B-10
Query Match 2.4%; Score 34; DB 3; Length 1582;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TITLE OF INVENTION: PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,071
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,166
FILING DATE: 27-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 95563/PS36321/US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1283 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-452-071-19
Query Match 2.4%; Score 34; DB 3; Length 1283;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1408 TTTTAAAAA 1441
DB 1244 TTTTAAAAA 1277
RESULT 14
US-09-614-912-191
Sequence 191, Application US/09614912
Patent No. 6677502
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Rafalski, Antoni
APPLICANT: Orozco, Buddy
APPLICANT: Miao, Gou-Hau
APPLICANT: Famodu, Omolayo O.
APPLICANT: Lee, Jian Ming
APPLICANT: Sakai, Hajime
APPLICANT: Weng, Zude
APPLICANT: Caimi, Perry G
APPLICANT: Anderson, Shawn
TITLE OF INVENTION: Plant Metabolism Genes
FILE REFERENCE: BB1378 US NA
CURRENT APPLICATION NUMBER: US/09/614,912
CURRENT FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: 60/143,401
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/143,412
PRIOR FILING DATE: 1999-07-12
PRIOR APPLICATION NUMBER: 60/146,650
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/170,906
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: 60/172,959
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/172,946
PRIOR FILING DATE: 1999-12-21
NUMBER OF SEQ ID NOS: 204
SOFTWARE: Microsoft Office 97
SEQ ID NO 191
LENGTH: 1518
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (114)
NAME/KEY: unsure
LOCATION: (123)
NAME/KEY: unsure
LOCATION: (138)
US-09-614-912-191
Query Match 2.4%; Score 34; DB 4; Length 1518;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1408 TTTTAAAAA 1441
DB 1434 TTTTAAAAA 1467
RESULT 15
US-08-545-196B-10
Sequence 10, Application US/08545196B
Patent No. 6080577
GENERAL INFORMATION:
APPLICANT: MELKI, JUDITH
APPLICANT: MUNNICH, ARNOLD
TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,196B
FILING DATE: 19-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FARACI, C. J.
REGISTRATION NUMBER: 32,350
REFERENCE/DOCKET NUMBER: 2121-110P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1582 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-545-196B-10
Query Match 2.4%; Score 34; DB 3; Length 1582;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1408 TTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1441
| | | | | | | | | | | | | | | | | | | | | |
Db 1485 TTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1518
| | | | | | | | | | | | | | | | | | | | | |

Search completed: June 3, 2004, 16:17:10
Job time : 126 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: June 3, 2004, 10:21:00 ; Search time 638 Seconds
(without alignments)
9595.068 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04; *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002s: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1441	100.0	1441	3	AAA94623 Human CAS
2	1413	98.1	1421	3	AAA96505 cDNA enco
3	1336	92.7	1491	9	ADD19228 Human cDN
4	1285	89.2	1474	6	ABL90535 Human pol
5	996	69.1	1460	3	AAC95564 Human sec
6	996	69.1	1460	7	ABZ67235 Human sec
7	996	69.1	1460	7	ABZ73640 Secreted
8	996	69.1	1460	9	ADC20289 Human sec
9	425	29.5	498	3	AAA94624 Human CAS
10	405	28.1	580	9	ADD19187 Human cDN
11	405	28.1	5033	7	ABZ68115 Human sec
12	405	28.1	5033	7	ABZ74587 Secreted
13	405	28.1	5033	9	ADC21005 Human sec
14	392	27.2	2684	7	ADA53627 Human cod
15	303	21.0	406	5	AAS66563 DNA encod
16	59	4.1	512	4	AAH36185 Human col
17	38	2.6	427	4	AAI87437 Human pol
18	38	2.6	999	6	ABN74679 Bovine em
19	38	2.6	2165	3	AAC90468 Human unc
20	38	2.6	5642	6	ABL32869 Human imm
21	38	2.6	10480	6	ABL34201 Human imm
22	36	2.5	896	4	AAH33168 Human col
23	36	2.5	2255	2	AAQ84698 Human IFN

C	24	35	2.4	297	5	ABV61920	Abv61920 Human pro
C	25	35	2.4	364	4	AAI93544	Aai93544 Human pro
C	26	35	2.4	436	5	ABV35678	Abv35678 Human pro
C	27	35	2.4	437	5	ABV44483	Abv44483 Human pro
C	28	35	2.4	450	6	ABZ08656	Abz08656 Human leu
C	29	35	2.4	453	4	AAI87364	Aai87364 Human pol
C	30	35	2.4	462	5	ABV34102	Abv34102 Human pro
C	31	35	2.4	462	5	ABV42965	Abv42965 Human pro
C	32	35	2.4	464	6	ABZ08188	Abz08188 Human leu
C	33	35	2.4	504	5	ABV56694	Abv56694 Human pro
C	34	35	2.4	572	5	ABV61258	Abv61258 Human pro
C	35	35	2.4	629	5	ABV12980	Abv12980 Human pro
C	36	35	2.4	644	4	AAI23651	Aai23651 Human bre
C	37	35	2.4	748	4	AAH71460	Aah71460 Human cer
C	38	35	2.4	830	5	ABV14600	Abv14600 Human pro
C	39	35	2.4	1034	3	AAZ52527	Aaz52527 Human sec
C	40	35	2.4	1639	5	AAZ85102	Aaz85102 Nucleotid
C	41	35	2.4	2857	9	ADD45824	Add45824 Human gen
C	42	35	2.4	2857	9	ADE60246	Ade60246 Human gen
C	43	35	2.4	2857	9	ADE60248	Ade60248 Human gen
C	44	35	2.4	2857	9	ADE60252	Ade60252 Human gen
C	45	35	2.4	2857	9	ADE60250	Ade60250 Human gen

ALIGNMENTS

RESULT 1
AAA94623
ID AAA94623 standard; DNA; 1441 BP.
XX
AC AAA94623;
XX
DT 11-JAN-2001 (first entry)
XX
DE Human CASB618 coding sequence.
XX
KW Spitope; human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian;
KW colon; autoimmune disease; HLA_A0201; ss.
XX
OS Homo sapiens.
XX
FH Key
CDS Location/Qualifiers
FT 259..1221
FT /*tag= a
FT /product= "Human CASB618"

XX
FN WO200053748-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-EP002048.
XX
PR 11-MAR-1999; 99GB-00005607.
PR 01-SEP-1999; 99GB-00020590.
XX
(SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Bruck CEM, Cassart J, Coche T, Vinals Y De BassolsC;
XX
WPI; 2000-572268/53.
DR P-PSDB; AAB26325.
XX
New human CASB618 polypeptide, useful as a vaccine for prophylactic and
therapeutic treatment of cancers, particularly ovarian or colon cancer,
autoimmune diseases and related conditions.
PS Claim 13; Page 61; 76pp; English.
XX
The present sequence is the coding sequence of human CASB618 protein. The
gene for human CASB618 is thought to be located on chromosome 15. The
protein encoded by the present sequence and epitopes of the CASB618
protein (see AAB26327 to AAB26399) are useful in diagnosing the occurrence

CC	of tumour cells and in vaccines for prophylactic and therapeutic
CC	treatment of cancers, particularly ovarian or colon cancer, autoimmune
CC	diseases and related conditions
XX	
SQ	Sequence 1441 BP; 289 A; 466 C; 392 G; 294 T; 0 U; 0 Other;
Query Match	100.0%; Score 1441; DB 3; Length 1441;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1441; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AAAGTAACGGGTACAGACAGTGGAGAAATAGTTTCGCTCGCCGGCTAGAAAACTCTGTCTG 60
DB	1 AAAGTAACGGGTACAGACAGTGGAGAAATAGTTTCGCTCGCCGGCTAGAAAACTCTGTCTG 60
QY	61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTTCAAGCTTTCCTTAACGGAGAGGTGCA 120
DB	61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTTCAAGCTTTCCTTAACGGAGAGGTGCA 120
QY	121 GGACTCAGACTTCACCCAGCCCTCGGTCCAGCTTGACGAAAGAGACGCCAAGGAC 180
DB	121 GGACTCAGACTTCACCCAGCCCTCGGTCCAGCTTGACGAAAGAGACGCCAAGGAC 180
QY	181 GCGCTCTCCCGCTCCAGGCAGCCAGCTTGCTGGCTTGCCCTGCCCGCTGGTGCGCAGC 240
DB	181 GCGCTCTCCCGCTCCAGGCAGCCAGCTTGCTGGCTTGCCCTGCCCGCTGGTGCGCAGC 240
QY	241 ACTCGGCCGGCTGAGCAGCATGACCTGTGGAAAGGCGTACTGCCCTTTTACCCCGAGCCC 300
DB	241 ACTCGGCCGGCTGAGCAGCATGACCTGTGGAAAGGCGTACTGCCCTTTTACCCCGAGCCC 300
QY	301 CCGCATGCGCAGGCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTGTTTGGCTCTA 360
DB	301 CCGCATGCGCAGGCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTGTTTGGCTCTA 360
QY	361 GCAGCAAGCTTCCTGCTCATCTTGC CGGGATCCGTGGCCACTCGCGCTGTTTGGTTG 420
DB	361 GCAGCAAGCTTCCTGCTCATCTTGC CGGGATCCGTGGCCACTCGCGCTGTTTGGTTG 420
QY	421 GTGAGAGTTCTTCTCAGTCTGTTTCATAGGGCGAGAAATTGTGGCTGTGCATTTCAGTGCA 480
DB	421 GTGAGAGTTCTTCTCAGTCTGTTTCATAGGGCGAGAAATTGTGGCTGTGCATTTCAGTGCA 480
QY	481 GAATGTTTCGTGGTACAGTGAAACACCAACATCTTACAAAGCCCTCAGCGCAGCGCGC 540
DB	481 GAATGTTTCGTGGTACAGTGAAACACCAACATCTTACAAAGCCCTCAGCGCAGCGCGC 540
QY	541 GTTACAGCCCGTGTGCTGTGCTGTGGGCTTGGAGGGCATTAATATTACATCACAGGG 600
DB	541 GTTACAGCCCGTGTGCTGTGCTGTGGGCTTGGAGGGCATTAATATTACATCACAGGG 600
QY	601 ACCCCAGTGCTACCTGAAACGAGACCATTTGACTACACGAGCAGTTTACCTGGCGTCTG 660
DB	601 ACCCCAGTGCTACCTGAAACGAGACCATTTGACTACACGAGCAGTTTACCTGGCGTCTG 660
QY	661 AAACAGAAATTACGCCCGGAGTACGCGAAACGCACTGGAGAGGGGCTGCCGAGCCAGTG 720
DB	661 AAACAGAAATTACGCCCGGAGTACGCGAAACGCACTGGAGAGGGGCTGCCGAGCCAGTG 720
QY	721 CTCTACTCTGGCGGAGAGTTTCAACCGAGTAGCCCTTGGCCCTGTACCCAGTACACAC 780
DB	721 CTCTACTCTGGCGGAGAGTTTCAACCGAGTAGCCCTTGGCCCTGTACCCAGTACACAC 780
QY	781 CTGGCGGACACTACGCTTCGGCCACGCTATGGTGGGCTTCTGCTTCTGGCTCTCTCC 840
DB	781 CTGGCGGACACTACGCTTCGGCCACGCTATGGTGGGCTTCTGCTTCTGGCTCTCTCC 840
QY	841 AACGTGCTGCTCTCCACGCCGGCCCGCTCTACGGAGGCTTGGCATGTGTGACACCGGA 900
DB	841 AACGTGCTGCTCTCCACGCCGGCCCGCTCTACGGAGGCTTGGCATGTGTGACACCGGA 900
QY	901 GCCTTCGCGCTCTTCGCGGTCTTCGCTTGGCTTCCATCTCTAGCGTCCGCTCTGCGCG 960
DB	901 GCCTTCGCGCTCTTCGCGGTCTTCGCTTGGCTTCCATCTCTAGCGTCCGCTCTGCGCG 960

XX WPI; 2000-579485/54.
DR P-PSDB; AAB18992.
XX New human transmembrane proteins are used to treat a disease or condition
PT associated with decreased expression of functional HTPP e.g. Tourette's
PT disorder, angina and leukemia.
XX Claim 4; Page 129; 130pp; English.
PS The present sequence encodes a human transmembrane proteins (HTMP).
XX Agonists and antagonists of the protein are used to treat a disease or
CC condition associated with overexpression of the protein. Diseases and
CC conditions which can be treated include cell proliferative,
CC immunological, reproductive, smooth muscle and neurological disorders
CC e.g. arteriosclerosis, myeloma, leukaemia, acquired immunodeficiency
CC syndrome (AIDS), allergies, ovulatory defects, angina, hypertension,
CC stroke, Alzheimer's disease, epilepsy and Tourette's disorder. The
CC polynucleotides may be used to detect and quantify gene expression in
CC biopsied tissues where protein expression may be correlated with disease
CC e.g. to determine absence, presence or excess expression of HTPP or to
CC monitor regulation of HTPP expression during therapeutic intervention
XX
SQ Sequence 1421 BP; 264 A; 468 C; 395 G; 294 T; 0 U; 0 Other;
Query Match 98.1%; Score 1413; DB 3; Length 1421;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGTAGAAAACCTGTGCG 60
DB 9 AAAGTAAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGTAGAAAACCTGTGCG 68
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCAGCTTCCTTAACGGAGAGTGCA 120
DB 69 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCAGCTTCCTTAACGGAGAGTGCA 128
QY 121 GGAATCAGACTTACCAAGCCCACTCGGTCCAGCTTGTACGCAAGAGACGCCAAGGAC 180
DB 129 GGAATCAGACTTACCAAGCCCACTCGGTCCAGCTTGTACGCAAGAGACGCCAAGGAC 188
QY 181 GGGCTCTCCCGCGTCCAGGCGCCAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
DB 189 GGGCTCTCCCGCGTCCAGGCGCCAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 248
QY 241 ACTGGCGCGCGTCCAGGCGCCAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
DB 249 ACTGGCGCGCGTCCAGGCGCCAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 308
QY 301 CGGATGCGCGCGTCCAGGCGCCAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360
DB 309 CGGATGCGCGCGTCCAGGCGCCAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 368
QY 361 GCAGCAAGCTTCTCTAGTCTGTTCATAGGCGCAGAAATGTGGCTGTGCACTTCAGTGCA 420
DB 369 GCAGCAAGCTTCTCTAGTCTGTTCATAGGCGCAGAAATGTGGCTGTGCACTTCAGTGCA 428
QY 421 GTGAGAGTTCTTCTAGTCTGTTCATAGGCGCAGAAATGTGGCTGTGCACTTCAGTGCA 480
DB 429 GTGAGAGTTCTTCTAGTCTGTTCATAGGCGCAGAAATGTGGCTGTGCACTTCAGTGCA 488
QY 481 GAATGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540
DB 489 GAATGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 548
QY 541 GTTACAGCCCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
DB 549 GTTACAGCCCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 608
QY 601 ACCCCAGTGCATCAGCTGAACGAGACCATTTGACTACACGAGCAGTTCACTGGCGTGTG 660
DB 609 ACCCCAGTGCATCAGCTGAACGAGACCATTTGACTACACGAGCAGTTCACTGGCGTGTG 668

RESULT 3
ADD19228
ID ADD19228 standard; cDNA; 1491 BP.
XX
AC ADD19228;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human cDNA from secreted protein gene 45.
XX human secreted protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal; cardiac;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW neutrotropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; human; ss; gene.

661 AAAGAGAATTACCGCGGAGTACCGAAGCACTGGAGAGGGGCTCCGAGCCAGTG 720
669 AAAGAGAATTACCGCGGAGTACCGAAGCACTGGAGAGGGGCTCCGAGCCAGTG 728
721 CTCTACCTGGCGGAGAAATTACACCGAGTAGCCCTTGGCGCTGTACACAGTACCAC 780
729 CTCTACCTGGCGGAGAAATTACACCGAGTAGCCCTTGGCGCTGTACACAGTACCAC 788
781 CTGGCGGAGACACTACCGCTCGGCCACGCTATGGGTGGCTTCTGCTTCTGCTCTCC 840
789 CTGGCGGAGACACTACCGCTCGGCCACGCTATGGGTGGCTTCTGCTTCTGCTCTCC 848
841 AACGTGCTGTCTCCACGCGCGCCCGCTCTACGAGGCGCTGGCTGTACACCGGA 900
849 AACGTGCTGTCTCCACGCGCGCGCCCGCTCTACGAGGCGCTGGCTGTACACCGGA 908
901 GCCTTCGCGCTCTTCGCGGTCTTCGCTTGGCTCCATCTCTAGCGTGGCTCTGCGCG 960
909 GCCTTCGCGCTCTTCGCGGTCTTCGCTTGGCTCCATCTCTAGCGTGGCTCTGCGCG 968
961 CTCCGCTAGGCTCTCCGCGCTCACCACTCAGTACGCGCGCGCTCTGCGGTACCGTG 1020
969 CTCCGCTAGGCTCTCCGCGCTCACCACTCAGTACGCGCGCGCTCTGCGGTACCGTG 1028
1021 GCAACCGCGCTCTTCGCTCTCTTCGAGGGGCGCTGGTGGTCTCCAGTATGTTCCG 1080
1029 GCAACCGCGCTCTTCGCTCTCTTCGAGGGGCGCTGGTGGTCTCCAGTATGTTCCG 1088
1081 CCCAGCGCTCTTCGACCCCTTCTGGACCAAGCGCCAGGACTGCAGCCAGGAGAGGG 1140
1089 CCCAGCGCTCTTCGACCCCTTCTGGACCAAGCGCCAGGACTGCAGCCAGGAGAGGG 1148
1141 GGCTCACTCTTATCTTCGCGGACCCCACTGCACAGCAGGCGCTCTCCAGACTTAAA 1200
1149 GGCTCACTCTTATCTTCGCGGACCCCACTGCACAGCAGGCGCTCTCCAGACTTAAA 1208
1201 TGTATCACCACTAACCTGTAGGGGGGACCAATCTGGACTCTTCCCGCGCTTGGGACAT 1260
1209 TGTATCACCACTAACCTGTAGGGGGGACCAATCTGGACTCTTCCCGCGCTTGGGACAT 1268
1261 CGCAGCCCGGAGAGCAGTCCCGCGCAGGCTGGGCGCAGAGAGCTCCAGGAGGGCAGTG 1320
1269 CGCAGCCCGGAGAGCAGTCCCGCGCAGGCTGGGCGCAGAGAGCTCCAGGAGGGCAGTG 1328
1321 AGCGTGTGGCGGAGGCTTCGGACATCCGCGAGGACCCAGGAAAGTCTCTCTGGGGCGA 1380
1329 AGCGTGTGGCGGAGGCTTCGGACATCCGCGAGGACCCAGGAAAGTCTCTCTGGGGCGA 1388
1381 TCTGTAATAAACCCTTTTCTTTTCTTTTCTTTT 1413
1389 TCTGTAATAAACCCTTTTCTTTTCTTTTCTTTT 1421

QY 1261 CGCAGCGCGGAGCACTGCGCCGCGCAGGCGCTGGCCAGGAGAGCTCCAGGAAGGGCACTG 1320
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1304 CGCAGCGCGGAGCACTGCGCCGCGCAGGCGCTGGCCAGGAGAGCTCCAGGAAGGGCACTG 1363
QY 1321 AGCGTCTGCGCGGAGGCGCTCGGACATCCGCGCAGGACACCGAGGAAAGTCTCTGGGCGGA 1380
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1364 AGCGTCTGCGCGGAGGCGCTCGGACATCCGCGCAGGACACCGAGGAAAGTCTCTGGGCGGA 1423
QY 1381 TCTGTAATAAACCTTTTCTTTTGTGTTTAAATAAAAAAAAAAAAAAAAAAAAA 1438
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1424 TCTGTAATAAACCTTTTCTTTTGTGTTTAAATAAAAAAAAAAAAAAAAAAAAA 1481

RESULT 4
ID ABL90535 standard; cDNA; 1474 BP.
XX
AC ABL90535;
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1097.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
FN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR P-PSDB; ABB90126.
XX

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 4; SEQ ID NO 1097; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB9040-ABB9044) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 1474 BP; 292 A; 480 C; 404 G; 296 T; 0 U; 2 Other;

Query Match 89.2%; Score 1285; DB 6; Length 1474;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1435; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCCGCTAGAAAACTCTGTGG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
27 AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCCGCTAGAAAACTCTGTGG 86
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCTTAACGGAGAGGTGCA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
87 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCTTAACGGAGAGGTGCA 146
QY 121 GGAATCACTTCAACAGCGCTTCCGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
147 GGAATCACTTCAACAGCGCTTCCGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 206
QY 181 GCGCTCTCCCGCTCCAGGAGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
207 GCGCTCTCCCGCTCCAGGAGCGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 266
QY 241 ACTCGGCGCGCTGCGAGCATGACCTGTGGAACCGCGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
267 ACTCGGCGCGCTGCGAGCATGACCTGTGGAACCGCGCTGCTGCTGCTGCTGCTGCTGCTG 326
QY 301 CGGCATCGCGCGCTTCCAGCGTTCCACTGCTCATCGTTATTTCTAGTGTGTTTGGCTCTA 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
327 CGGCATCGCGCGCTTCCAGCGTTCCACTGCTCATCGTTATTTCTAGTGTGTTTGGCTCTA 386
QY 361 GCAGCAAGCTTCTGCTCATCTTCCGCGGGATCGGTGGCCACTCGCGCTGCTGTTGTTG 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
387 GCAGCAAGCTTCTGCTCATCTTCCGCGGGATCGGTGGCCACTCGCGCTGCTGTTGTTG 446
QY 421 GTGAGAGTTCTTCTCAGTCTGTTTATAGCGCGCAAGAAATTTGTGGCTGTGCACTTCA 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
447 GTGAGAGTTCTTCTCAGTCTGTTTATAGCGCGCAAGAAATTTGTGGCTGTGCACTTCA 506
QY 481 GAATGTTTCTGTTGAGTGAACCAACACATCTCTACAAAGCCTTCAGCGCAGCGCGC 540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
507 GAATGTTTCTGTTGAGTGAACCAACACATCTCTACAAAGCCTTCAGCGCAGCGCGC 566
QY 541 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
567 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
QY 601 ACCCAGTGCATCAGTGAACGAGACCATTTGACTACACGAGCAGTTTACCTGGGCTCTG 660
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
627 ACCCAGTGCATCAGTGAACGAGACCATTTGACTACACGAGCAGTTTACCTGGGCTCTG 686
QY 661 AAAGAGAATTACCGCGCGAGTACCGGAACGCACTGGAGAAGGGGCTGCCGACCCAGTG 720
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
687 AAAGAGAATTACCGCGCGAGTACCGGAACGCACTGGAGAAGGGGCTGCCGACCCAGTG 746
QY 721 CTCTACTGGCGGAGAGTTTACACCGAGTAGCCCTTGGGCTGTACACACAGTACCCAC 780
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
747 CTCTACTGGCGGAGAGTTTACACCGAGTAGCCCTTGGGCTGTACACACAGTACCCAC 806
QY 781 CTGGCGGACACTACGCTCGGCGCAGCTATGGGTGGGTTCTGCTTCTGGGCTCTCTCTCC 840
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
807 CTGGCGGACACTACGCTCGGCGCAGCTATGGGTGGGTTCTGCTTCTGGGCTCTCTCTCC 866
QY 841 AACGTGCTCTCTCCAGCGCGCGCGCTCTACCGAGGCTGCGCACTGTGACCAACCGGA 900
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
867 AACGTGCTCTCTCCAGCGCGCGCGCTCTACCGAGGCTGCGCACTGTGACCAACCGGA 926
QY 901 GCCTTCGCGCTCTTCCGCGCTTCTGCGCTTGGCTCCACTCTAGCGTGGGCTCTGCGCG 960
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
927 GCCTTCGCGCTCTTCCGCGCTTCTGCGCTTGGCTCCACTCTAGCGTGGGCTCTGCGCG 986
QY 961 CTCCGCTTAGGCTCTCTCCGCGCTCACCACTCAGTACGGCGCGCGCTTCTGGGTCACGCTG 1020
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
987 CTCCGCTTAGGCTCTCTCCGCGCTCACCACTCAGTACGGCGCGCGCTTCTGGGTCACGCTG 1046
QY 1021 GCAACCGCGCTCTCTGCTGCTTCTTCTCGAGGGGCGCGGTGAGTCTCCAGTATGTTGG 1080

Db 682 AAGAGAAATTACGCCGCGGAGTAGCGAAGCACTGGAGAGGGGCTGCCGACCCAGTG 741
Qy 721 CTCTACCTGGGAGAGATTCAACCGAGTAGCCCTTGGGCTGTACCACAGTACCAC 780
Db 742 CTCTACCTGGGAGAGATTCAACCGAGTAGCCCTTGGGCTGTACCACAGTACCAC 801
Qy 781 CTGGCGGGACACTAGCCTCGGCGACGCTATGGGTGGCTTCTGCTTCTGCTCTCTCC 840
Db 802 CTGGCGGGACACTAGCCTCGGCGACGCTATGGGTGGCTTCTGCTTCTGCTCTCTCC 861
Qy 841 AACGTGCTGCTCTCCACCGCGGCGGCTCTACGGAGGCGCTGGCACTGTGACCAACCGGA 900
Db 862 AACGTGCTGCTCTCCACCGCGGCGGCTCTACGGAGGCGCTGGCACTGTGACCAACCGGA 921
Qy 901 GCCTTCGGCTCTTCGGGGCTCTTCGGCTTGGCTCCATCTCTAGCGTGCCTGCTGCGCG 960
Db 922 GCCTTCGGCTCTTCGGGGCTCTTCGGCTTGGCTCCATCTCTAGCGTGCCTGCTGCGCG 981
Qy 961 CTCGGCTAGGCTCTCCGCGCTCACCACCTCAGTACGGCGCGCTTCTGGGTCAAGCTG 1020
Db 982 CTCGGCTAGGCTCTCCGCGCTCACCACCTCAGTACGGCGCGCTTCTGGGTCAAGCTG 1041
Qy 1021 GCAACCGGCGTCTGTGCTCTTCCTC 1047
Db 1042 GCAACCGGCGTCTGTGCTCTTCCTC 1068

RESULT 6

ABZ67235
ID ABZ67235 standard; cDNA; 1460 BP.
XX
AC ABZ67235;
XX
DT 26-MAR-2003 (first entry)
XX
DE Human secreted protein encoding cDNA SEQ ID NO 355.
XX
KW Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnary; antibacterial; antiparkinsonian; antischistosomal; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; chromosome 9p21; ds.
XX
OS Homo sapiens.
XX
PN WO200277186-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-US009188.
XX
PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-040583/03.
DR P-PSDB; ABP99814.
XX
PT New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGCAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX
PS Claim 7; Page 1353; 2423pp; English.
XX
CC The invention relates to novel human genes (ABZ66891-ABZ68209) and the

CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (antagonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
XX
SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;
Query Match 69.1%; Score 996; DB 7; Length 1460;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1046; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AAAGTAACGGCTACAGACAGTGAATAAGTTTTCGTCGCGGCTAGAAAACTCTGTGC 60
Db 22 AAAGTAACGGCTACAGACAGTGAATAAGTTTTCGTCGCGGCTAGAAAACTCTGTGC 81
Qy 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAAACGAGAGGTGCA 120
Db 82 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAAACGAGAGGTGCA 141
Qy 121 GGACTCAGACTTCAACGACCCACTCGGTCCAGCCCTGTACGCAAGAGAGCCAGGAC 180
Db 142 GGACTCAGACTTCAACGACCCACTCGGTCCAGCCCTGTACGCAAGAGAGGTGCA 201
Qy 181 GCGCTCTCCGCGTCCAGGAGCCCGAGCTTGTGGCTTGGCTGCCGCTGCGTGCAGC 240
Db 202 GCGCTCTCCGCGTCCAGGAGCCCGAGCTTGTGGCTTGGCTGCCGCTGCGTGCAGC 261
Qy 241 ACTCGGCGCGGCTGCGATGACCCCTGTGGAACGGCTGCTTCTTACCCCGAGCC 300
Db 262 ACTCGGCGCGGCTGCGATGACCCCTGTGGAACGGCTGCTTCTTACCCCGAGCC 321
Qy 301 GGGCATGCGCGAGGCTTCCAGCTTCCAGCTCATCGTTATTCTAGTGTGTTTGGCTCTA 360
Db 322 GGGCATGCGCGAGGCTTCCAGCTTCCAGCTCATCGTTATTCTAGTGTGTTTGGCTCTA 381
Qy 361 GCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGGCACTCGCGCTGTTGTTG 420
Db 382 GCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGGCACTCGCGCTGTTGTTG 441
Qy 421 GTGAGAGTCTTCTCAGTCTGTTCATAGGCGCAGAAATGTGGTGTGCACTTCAAGTGA 480
Db 442 GTGAGAGTCTTCTCAGTCTGTTCATAGGCGCAGAAATGTGGTGTGCACTTCAAGTGA 501
Qy 481 GAATGTTGTTGGTACAGTGAACACCAACACATCTTACAAAGCCCTTACGCGCAGCGGC 540
Db 502 GAATGTTGTTGGTACAGTGAACACCAACACATCTTACAAAGCCCTTACGCGCAGCGGC 561
Qy 541 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 562 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
Qy 601 ACCCCAGTGCATCAGCTGAACGAGACCAATGACTACAGAGCAGTTTACCTGGCGTCTG 660
Db 622 ACCCCAGTGCATCAGCTGAACGAGACCAATGACTACAGAGCAGTTTACCTGGCGTCTG 681
Qy 661 AAAGAGAAATTACCGCGGAGTACCGGACGCACTGGAGAGGGGCTGCCGACCCAGTG 720
Db 682 AAAGAGAAATTACCGCGGAGTACCGGACGCACTGGAGAGGGGCTGCCGACCCAGTG 741
Qy 721 CTCTACCTGGCGGAGAGTTTCAACCGAGTAGCCCTTGGCGCTGTACCAACAGTACCAC 780
Db 742 CTCTACCTGGCGGAGAGTTTCAACCGAGTAGCCCTTGGCGCTGTACCAACAGTACCAC 801

QY 781 CTGGCGGACACTAGCCTCGGTCACGCTATGGTGGGTTCTGCTTCTGCTCTCTCC 840
Db |||||
QY 802 CTGGCGGACACTAGCCTCGGTCACGCTATGGTGGGTTCTGCTTCTGCTCTCTCC 861
Db |||||
QY 841 AACGTGCTCTCTCCACGCGCGCCCGCTCTACGAGGCGCTGCACTGCTGACCAACCGGA 900
Db |||||
QY 862 AACGTGCTCTCTCCACGCGCGCCCGCTCTACGAGGCGCTGCACTGCTGACCAACCGGA 921
Db |||||
QY 901 GCCTTCGGCTCTCTGGGGTCTTCGCTTGGCTTCCATCTCTAGCGTGCCTGCTGCGCG 960
Db |||||
QY 922 GCCTTCGGCTCTCTGGGGTCTTCGCTTGGCTTCCATCTCTAGCGTGCCTGCTGCGCG 981
Db |||||
QY 961 CTCGCTAGGCTCTCTCGGCTCTCCACCACTCAGTACGCGCGCTTCTGCTGCTGCTG 1020
Db |||||
QY 982 CTCGCTAGGCTCTCTCGGCTCTCCACCACTCAGTACGCGCGCTTCTGCTGCTGCTG 1041
Db |||||
QY 1021 GCAACCGGCTCTCTGCT 1047
Db |||||
QY 1042 GCAACCGGCTCTCTGCT 1068
Db |||||
RESULT 7
ABZ73640
ID ABZ73640 standard; cDNA; 1460 BP.
XX
AC ABZ73640;
XX
XX
DT 12-MAY-2003 (first entry)
XX
DS Secreted protein-encoding gene 360 cDNA clone HUFCL31, SEQ ID NO:370.
XX
KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW drug screening; chromosome identification; chromosome mapping;
KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW antianaemic; vulnery; chromosome 9p21; gene; ss.
XX
OS Homo sapiens.
XX
XX
PN WO200277013-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-US009370.
XX
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
PI
XX
XX WPI; 2003-040578/03.
DR P-PSDB; ABR01306.
DR
XX
PT New human secreted proteins and nucleic acids, useful for detecting or
PT treating cancer or other hyperproliferative disorders, autoimmune
PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
XX
PS Claim 21; Page 1345; 2474pp; English.
XX
XX ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins are thought to be involved in biological activities
CC associated with cellular signalling, cellular differentiation, cell
CC migration, prohormone activation and neurotransmitter activity. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody

CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing or treating cancers or other
CC hyperproliferative disorders. Additionally, the secreted proteins and
CC their nucleic acids may also be used in the treatment of autoimmune
CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC wound healing. Nucleic acids of the invention may be used for chromosome
CC identification, chromosome mapping, in gene therapy, for identifying
CC individuals from minute biological samples, as hybridisation probes, and
CC as molecular weight markers. The present sequence represents a human
CC secreted protein-encoding cDNA clone of the invention
XX
SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;
Query Match 69.1%; Score 996; DB 7; Length 1460;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1046; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTCGCGCTAGAAAAAATCTGTGCG 60
Db |||||
QY 22 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTCGCGCTAGAAAAAATCTGTGCG 81
Db |||||
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAACGAGAGGTGCA 120
Db |||||
QY 82 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAACGAGAGGTGCA 141
Db |||||
QY 121 GGACTCAGACTTCAACAGCCACTCGGTCCAGAGCCTTGTACGCAAAAGAGAGCGCAAGGAC 180
Db |||||
QY 142 GGACTCAGACTTCAACAGCCACTCGGTCCAGAGCCTTGTACGCAAAAGAGAGCGTCAAGGAC 201
Db |||||
QY 181 GCGCTCTCCGCGTCCAGAGCCTCCAGCTTGTGGCTTGGCTGCGCGCTGCGTGCAGC 240
Db |||||
QY 202 GCGCTCTCCGCGTCCAGAGCCTCCAGCTTGTGGCTTGGCTGCGCGCTGCGTGCAGC 261
Db |||||
QY 241 ACTCGCGCGGCTGAGCATGACCTCTGTGGAGCGCGTACTGCTTTTACCCCGAGCCC 300
Db |||||
QY 262 ACTCGCGCGGCTGAGCATGACCTCTGTGGAGCGCGTACTGCTTTTACCCCGAGCCC 321
Db |||||
QY 301 CGGCATGCGCGAGGCTTCAAGCTTCCATCTGCTCATCGTTATTTCTAGTGTGTTGGCTCTA 360
Db |||||
QY 322 CGGCATGCGCGAGGCTTCAAGCTTCCATCTGCTCATCGTTATTTCTAGTGTGTTGGCTCTA 381
Db |||||
QY 361 GCAGCAAGCTTCTGCTCATCTGCTCGCGGATCCGTCGCGCTGCTGTTGGTTG 420
Db |||||
QY 382 GCAGCAAGCTTCTGCTCATCTGCTCGCGGATCCGTCGCGCTGCTGTTGGTTG 441
Db |||||
QY 421 GTGAGAGTTCTTCTCAGTCTGTTTATAGCGCAGAAATTTGGCTGTGCATTCAGTGCA 480
Db |||||
QY 442 GTGAGAGTTCTTCTCAGTCTGTTTATAGCGCAGAAATTTGGCTGTGCATTCAGTGCA 501
Db |||||
QY 481 GAATGGTTCTGTTGTTACAGTGAACCAACACATCTTACAAAGCCTTTCAGCGCAGCGCGC 540
Db |||||
QY 502 GAATGGTTCTGTTGTTACAGTGAACCAACACATCTTACAAAGCCTTTCAGCGCAGCGCGC 561
Db |||||
QY 541 GTTACAGCCCGTCTGCTGCTGCTGCGGCTGAGGCGCATTAATATACACTCACAGGG 600
Db |||||
QY 562 GTTACAGCCCGTCTGCTGCTGCTGCGGCTGAGGCGCATTAATATACACTCACAGGG 621
Db |||||
QY 601 ACCCAGTGCATCAGCTGAACGAGACCATTTGACTACAAAGCAGTTCACCTGGGCTCTG 660
Db |||||
QY 622 ACCCAGTGCATCAGCTGAACGAGACCATTTGACTACAAAGCAGTTCACCTGGGCTCTG 681
Db |||||
QY 661 AAAGAGATTACGCGCGGAGTACGCGAAGCAGTGAAGGAGGCTGCCGACCCAGTG 720
Db |||||
QY 682 AAAGAGATTACGCGCGGAGTACGCGAAGCAGTGAAGGAGGCTGCCGACCCAGTG 741
Db |||||
QY 721 CTCTACCTGGCGGAGAGTTTCAACCGAGTAGCCCTTGGCGCTGTACCAACAGTACAC 780
Db |||||
QY 742 CTCTACCTGGCGGAGAGTTTCAACCGAGTAGCCCTTGGCGCTGTACCAACAGTACAC 801
Db |||||
QY 781 CTGCGGGGACACTACGCTCGGCGACGCTATGGTGGGCTTCTGCTTCTGGCTCCTCTCC 840
Db |||||
QY 802 CTGCGGGGACACTACGCTCGGCGACGCTATGGTGGGCTTCTGCTTCTGGCTCCTCTCC 861
Db |||||

QY 841 AACGTGCTGCTTCCACGGCGGCCCGCTCTACGAGGCGCTGGCACTGCTGACCAACCGGA 900
DB |||||
862 AACGTGCTGCTTCCACGGCGGCCCGCTCTACGAGGCGCTGGCACTGCTGACCAACCGGA 921
QY 901 GCCTTCGGCTCTTCGGGGCTCTTCGGCTTGGCTCCATCTCTAGAGTGGCGCTCTGCGCG 960
DB |||||
922 GCCTTCGGCTCTTCGGGGCTCTTCGGCTTGGCTCCATCTCTAGAGTGGCGCTCTGCGCG 981
QY 961 CTCGGCTAGGCTCTTCGGGGCTTCACCACTCAGTACGGCGCGCTTCCTGGGTACGCTG 1020
DB |||||
982 CTCGGCTAGGCTCTTCGGGGCTTCACCACTCAGTACGGCGCGCTTCCTGGGTACGCTG 1041
QY 1021 GCAACCGGCGTCTGCTGCTCTTCCTC 1047
DB |||||
1042 GCAACCGGCGTCTGCTGCTCTTCCTC 1068

RESULT 8

ID ADC20289 standard; DNA; 1460 BP.
XX
AC ADC20289;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human secreted protein coding sequence #228.
XX
KW gene therapy; human; secreted protein; haemopoietic disorder;
KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
KW leukaemia; wound healing; epithelial cell proliferation disorder;
KW immune disorder; autoimmune disorder; asthmatic disorder;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.

XX Homo sapiens.
OS
XX
XX WO200292787-A2.
XX
PD 21-NOV-2002.
XX
PF 26-MAR-2002; 2002WO-US009257.
XX
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-129287/12.
DR
XX
PT New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating hematopoietic or hematologic disorders, e.g.
PT anemia or hemophilia.
XX
PS Claim 1; SEQ ID NO 238; 1512pp; English.

XX
CC The invention comprises the amino acid and coding sequences of human
CC secreted proteins. The DNA and protein sequences of the invention are
CC useful for detecting, preventing, diagnosing, prognosticating, treating
CC or ameliorating: haematopoietic or haematological disorders (e.g. anaemia
CC and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
CC and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
CC wound healing and disorders of epithelial cell proliferation; immune
CC disorders (e.g. autoimmune disorders and asthmatic disorders);
CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC and gastrointestinal disorders (e.g. duodenal ulcers and

CC gastroenteritis). The present DNA sequence encodes a human secreted
CC protein of the invention.
XX
SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;
Query Match 69.1%; Score 996; DB 9; Length 1460;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1046; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAGTAACGGCTACAGACAGTGAATAAGTTTCGCTGCGCGGCTAGAAAACTCTGTG 60
DB |||||
22 AAAGTAACGGCTACAGACAGTGAATAAGTTTCGCTGCGCGGCTAGAAAACTCTGTG 81
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAACGGAGAGGTGCA 120
DB |||||
82 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAACGGAGAGGTGCA 141
QY 121 GGACTCAGACTTCAACGACCCACTCGGTCCAGCCCTTGTACGCAAGAGACGCCAAGGAC 180
DB |||||
142 GGACTCAGACTTCAACGACCCACTCGGTCCAGCCCTTGTACGCAAGAGACGTCAGGAC 201
QY 181 GCGCTCTCCGCGTCCAGGAGCCCGCTTGTGGCTTGGCTGCGCGCTGCGTGCAGC 240
DB |||||
202 GCGCTCTCCGCGTCCAGGAGCCCGCTTGTGGCTTGGCTGCGCGCTGCGTGCAGC 261
QY 241 ACTCGGCGCGGTGCAGATGACCTGTGGAACGGGTACTGCTTTTACCCCGCAGCC 300
DB |||||
262 ACTCGGCGCGGTGCAGATGACCTGTGGAACGGGTACTGCTTTTACCCCGCAGCC 321
QY 301 CGGCATGCGCGCAGGCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTTTTGGCTCTA 360
DB |||||
322 CGGCATGCGCGCAGGCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTTTTGGCTCTA 381
QY 361 GCAGCAAGCTTCTGCTCATCTTCCCGGGATCCGTGGCACTCGCGCTGGTTGGTTG 420
DB |||||
382 GCAGCAAGCTTCTGCTCATCTTCCCGGGATCCGTGGCACTCGCGCTGGTTGGTTG 441
QY 421 GTGAGAGTTCTTCTCAGTCTGTTTATAGGCGCAGAAATTTGGTGTGTGCACTTCAGTGCA 480
DB |||||
442 GTGAGAGTTCTTCTCAGTCTGTTTATAGGCGCAGAAATTTGGTGTGTGCACTTCAGTGCA 501
QY 481 GAATGGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 540
DB |||||
502 GAATGGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 561
QY 541 GTTACAGCCCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 600
DB |||||
562 GTTACAGCCCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 621
QY 601 ACCCCAGTGCATCAGCTGAACGAGACCAATTGACTACAACGAGAGTTTCACTGGCGCTG 660
DB |||||
622 ACCCCAGTGCATCAGCTGAACGAGACCAATTGACTACAACGAGAGTTTCACTGGCGCTG 681
QY 661 AAAGAGAAATTACGCGCGGAGTACCGAAACGCACTGGAGAGGGGCTGCCGAGCCCACTG 720
DB |||||
682 AAAGAGAAATTACGCGCGGAGTACCGAAACGCACTGGAGAGGGGCTGCCGAGCCCACTG 741
QY 721 CTCTACCTGGCGGAGAAAGTTCAACCGAGTACCGCTTGGCGCTGTACCAACAGTACCAC 780
DB |||||
742 CTCTACCTGGCGGAGAAAGTTCAACCGAGTACCGCTTGGCGCTGTACCAACAGTACCAC 801
QY 781 CTGGCGGGACACTACGCTTGGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 840
DB |||||
802 CTGGCGGGACACTACGCTTGGCGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTT 861
QY 841 AACGTGCTGCTTCCACGGCGGCCCGCTCTACGAGGCGCTGGCACTGCTGACCAACCGGA 900
DB |||||
862 AACGTGCTGCTTCCACGGCGGCCCGCTCTACGAGGCGCTGGCACTGCTGACCAACCGGA 921
QY 901 GCCTTCGGCTCTTCGGGGCTCTTCGGCTTGGCTCCATCTCTAGAGTGGCGCTCTGCGCG 960
DB |||||
922 GCCTTCGGCTCTTCGGGGCTCTTCGGCTTGGCTCCATCTCTAGAGTGGCGCTCTGCGCG 981

QY 961 CTCGGCTAGGCTCCTCGCGCTCACCACTCAGTACGGCGCGCCTTCTGGGTACGCTG 1020
Dd |||||
QY 982 CTCGGCTAGGCTCCTCGCGCTCACCACTCAGTACGGCGCGCCTTCTGGGTACGCTG 1041
Dd |||||
QY 1021 GCAACGGCGCTCCTGTGCTCTTCCTC 1047
Dd |||||
1042 GCAACGGCGCTCCTGTGCTCTTCCTC 1068
Dd |||||
RESULT 9
ID AAA94624 standard; DNA; 498 BP.
XX
AC AAA94624;
XX
DT 11-JAN-2001 (first entry)
XX
DE Human CASB618 EST.
XX
KW Epitope; human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian;
KW colon; autoimmune disease; HLA_A0201; expressed sequence tag; EST; ss.
XX
OS Homo sapiens.
XX
PN WO200053748-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-BP002048.
XX
PR 11-MAR-1999; 99GB-00005607.
PR 01-SEP-1999; 99GB-00020590.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Bruck CEM, Cassart J, Coche T, Vinals Y De BassolsC;
XX
DR WPI; 2000-572268/53.
XX
PT New human CASB618 polypeptide, useful as a vaccine for prophylactic and
PT therapeutic treatment of cancers, particularly ovarian or colon cancer,
PT autoimmune diseases and related conditions.
XX
PS Claim 32; Page 62; 76pp; English.
XX
CC The present sequence is an expressed sequence tag (EST) for human CASB618
CC protein. The gene for human CASB618 is thought to be located on
CC chromosome 15. CASB618 protein and epitopes of CASB618 protein (see
CC AAB26327 to AAB26399) are useful in diagnosing the occurrence of tumour
CC cells and in vaccines for prophylactic and therapeutic treatment of
CC cancers, particularly ovarian or colon cancer, autoimmune diseases and
CC related conditions
XX
SQ Sequence 498 BP; 97 A; 166 C; 137 G; 98 T; 0 U; 0 Other;
Query Match 29.5%; Score 425; DB 3; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.9e-153;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 998 GCGCGGCTTCTGGGTACCGTGGCAACCGCGCTCCTGTGCTCTTCTCGGAGGGCGG 1057
Dd |||||
62 GCGCGGCTTCTGGGTACCGTGGCAACCGCGCTCCTGTGCTCTTCTCGGAGGGCGG 121
QY 1058 TGGTGAGTCTCCAGTATGTTGGGCCAGCGCTCTTCGACCCCTTCTGGACCAAGCGCCA 1117
Dd |||||
122 TGGTGAGTCTCCAGTATGTTGGGCCAGCGCTCTTCGACCCCTTCTGGACCAAGCGCCA 181
QY 1118 AGGACTGCAGCCAGGAGAGAGGGGGCTCACCTTATCTTCGGCGACCCACTGCACAAGC 1177
Dd |||||
182 AGGACTGCAGCCAGGAGAGAGGGGGCTCACCTTATCTTCGGCGACCCACTGCACAAGC 241
QY 1178 AGGCGGCTTCTCCAGACTTAAATGTATCACTAACCCTGTGAGGGGAGCCCAATCTGG 1237
Dd |||||

Dd 242 AGGCGGCTTCTCCAGACTTAAATGTATCACTAACCCTGTGAGGGGAGCCCAATCTGG 301
QY 1238 ACTCCTTCCCGCCTTGGGACATCGCAGGCCGGGAAGCAGTCCCGCCAGGCTGGGCCA 1297
Dd |||||
302 ACTCCTTCCCGCCTTGGGACATCGCAGGCCGGGAAGCAGTCCCGCCAGGCTGGGCCA 361
QY 1298 GGAGAGCTCCAGGAAGGGCACTGAGCGCTGTGGCGCGAGGCTCGGACATCCGAGGCA 1357
Dd |||||
362 GGAGAGCTCCAGGAAGGGCACTGAGCGCTGTGGCGCGAGGCTCGGACATCCGAGGCA 421
QY 1358 CCAGGGAAAGTCTCCTGGGGGATCTGTAAATAAACCTTTTCTTTTGTTTTAAAA 1417
Dd |||||
422 CCAGGGAAAGTCTCCTGGGGGATCTGTAAATAAACCTTTTCTTTTGTTTTAAAA 481
QY 1418 AAAAA 1422
Dd |||||
482 AAAAA 486
RESULT 10
ID ADD19187 standard; cDNA; 580 BP.
XX
AC ADD19187;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human cDNA from secreted protein gene 4.
XX
KW human secreted protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiac;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW neutropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; human; ss; gene.
XX
OS Homo sapiens.
PN WC2003052377-A2.
XX
PD 26-JUN-2003.
XX
PF 06-NOV-2002; 2002WO-US035606.
XX
PR 07-NOV-2001; 2001US-0331046P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-533050/50.
DR P-PSDB; ADD19262.
XX
PT New isolated nucleic acids encoding signal transduction pathway component
PT polypeptides, useful for diagnosing, treating, and/or preventing
PT disorders, such as cancer, infections, cardiovascular and inflammatory
PT diseases.
XX
PS Claim 1; SEQ ID NO 14; 554pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human secreted protein, representing one of 85 novel genes.
CC Also included are recombinant vectors, host cells (expressing the
CC protein), the secreted proteins (including their fragments, epitopes and
CC homologues), an isolated antibody that binds specifically to the protein,
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition (comprising determining the presence or absence of a mutation
CC in the nucleic acid and diagnosing a condition based on the presence or
CC absence of the mutation), diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the

CC presence or amount of expression of the protein in a biological sample
CC and diagnosing a condition based on the presence or amount of expression
CC of the protein), preventing, treating or ameliorating a medical condition
CC by administering the nucleic acid or protein to a mammalian subject,
CC identifying a binding partner to the protein, the gene corresponding to
CC the cDNA sequence, and identifying an activity in a biological assay
CC (comprising expressing the nucleic acid in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity). The nucleic acids
CC and proteins display the following activities Cytostatic, antibacterial,
CC Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen,
CC Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-
CC Gen, Respiratory-Gen, Immunosuppressive, Cerebroprotective, Vasotropic,
CC Nootropic, Antiallergic. The methods and compositions of the present
CC invention are useful for diagnosing, treating, preventing and/or
CC prognosticating disorders related to the novel polypeptides, such as
CC cancer, bacterial or viral infections, and neural, immune system, blood,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, inflammatory or proliferative disorders (many examples of these
CC diseases and disorders are given in the specification). The present
CC sequence encodes a novel secreted protein of the invention.
XX
SQ Sequence 580 BP; 115 A; 198 C; 156 G; 111 T; 0 U; 0 Other;

Query Match 28.1%; Score 405; DB 9; Length 580;
Best Local Similarity 100.0%; Pred. No. 8.6e-146;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAAACGGCTACAGACAGTGCAGAAATAGTTTCGCTCGCGCTAGAAAACCTGTGTCG 60
DB 13 AAAGTAAACGGCTACAGACAGTGCAGAAATAGTTTCGCTCGCGCTAGAAAACCTGTGTCG 72
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCCTTAACGGAGAGGTGCA 120
DB 73 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCCTTAACGGAGAGGTGCA 132
QY 121 GCACTCAGACTTCACAGCCCACTCGGTCCAGCCCTTGTAACGAAAGAGACGCCAAGGAC 180
DB 133 GCACTCAGACTTCACAGCCCACTCGGTCCAGCCCTTGTAACGAAAGAGACGCCAAGGAC 192
QY 181 GCGCTCTCCCGGCTCCAGGAGCCCACTCGGTCCAGCCCTTGTAACGAAAGAGAGCGTGCAGC 240
DB 193 GCGCTCTCCCGGCTCCAGGAGCCCACTCGGTCCAGCCCTTGTAACGAAAGAGAGCGTGCAGC 252
QY 241 ACTCGGCGCGGTGTCAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCAAGCC 300
DB 253 ACTCGGCGCGGTGTCAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCAAGCC 312
QY 301 CGGATGCGCGAGCTTCAGCGTTCCACTGCTCATCGTTATCTAGTGTGTTTGGCTCTA 360
DB 313 CGGATGCGCGAGCTTCAGCGTTCCACTGCTCATCGTTATCTAGTGTGTTTGGCTCTA 372
QY 361 GCAGCAAGCTTCTGCTCATCTTCGCGGGGATCCGTGGCCACTCG 405
DB 373 GCAGCAAGCTTCTGCTCATCTTCGCGGGGATCCGTGGCCACTCG 417

RESULT 11
ABZ68115
ID ABZ68115 standard; DNA; 5033 BP.
XX
AC ABZ68115;
XX
XX 26-MAR-2003 (first entry)
DT
XX
DE Human secreted protein encoding genomic DNA SEQ ID NO 1638.
XX
KW Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnary; antibacterial; antiparkinsonian; antiskinning; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiac; immune disorder; infection; vaccine;

KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.
XX Homo sapiens.
XX
PN WO200277186-A2.
XX
PD 03-OCT-2002.
XX
XX 26-MAR-2002; 2002WO-US009188.
PF
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI WPI; 2003-040583/03.
XX
PT New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGCAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX
PS Disclosure; Page 2263-2264; 2423pp; English.
XX

CC The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (antagonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis; rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
XX
SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match 28.1%; Score 405; DB 7; Length 5033;
Best Local Similarity 100.0%; Pred. No. 6e-146;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAAACGGCTACAGACAGTGCAGAAATAGTTTCGCTCGCGCTAGAAAACCTGTGTCG 60
DB 13 AAAGTAAACGGCTACAGACAGTGCAGAAATAGTTTCGCTCGCGCTAGAAAACCTGTGTCG 72
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCCTTAACGGAGAGGTGCA 120
DB 73 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCCTTAACGGAGAGGTGCA 132
QY 121 GCACTCAGACTTCACAGCCCACTCGGTCCAGCCCTTGTAACGAAAGAGACGCCAAGGAC 180
DB 133 GCACTCAGACTTCACAGCCCACTCGGTCCAGCCCTTGTAACGAAAGAGACGCCAAGGAC 192
QY 181 GCGCTCTCCCGGCTCCAGGAGCCCACTCGGTCCAGCCCTTGTAACGAAAGAGAGCGTGCAGC 240
DB 193 GCGCTCTCCCGGCTCCAGGAGCCCACTCGGTCCAGCCCTTGTAACGAAAGAGAGCGTGCAGC 252
QY 241 ACTCGGCGCGGTGTCAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCAAGCC 300
DB 253 ACTCGGCGCGGTGTCAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCAAGCC 312
QY 301 CGGATGCGCGAGCTTCAGCGTTCCACTGCTCATCGTTATCTAGTGTGTTTGGCTCTA 360
DB 313 CGGATGCGCGAGCTTCAGCGTTCCACTGCTCATCGTTATCTAGTGTGTTTGGCTCTA 372

QY 361 GCAGCAAGCTTCCTGCTCATCTTGGCGGGATCCGTCGCACTCG 405
 DB 373 GCAGCAAGCTTCCTGCTCATCTTGGCGGGATCCGTCGCACTCG 417

RESULT 12

ABZ74587
 ID ABZ74587 standard; DNA; 5033 BP.

XX AC ABZ74587;

DT 12-MAY-2003 (first entry)

DE Secreted protein gene 360 genomic fragment HUFCL31, SEQ ID NO:1734.

XX Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KW drug screening; chromosome identification; chromosome mapping;
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 KW antianaemic; vulnery; chromosome 9p21; gene; ds.

OS Homo sapiens.

XX WO200277013-A2.

XX 03-OCT-2002.

XX 26-MAR-2002; 2002WO-US009370.

XX 27-MAR-2001; 2001US-0278650P.

XX 12-SEP-2001; 2001US-00950082.

XX 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-040578/93.

XX New human secreted proteins and nucleic acids, useful for detecting or
 PT treating cancer or other hyperproliferative disorders, autoimmune
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
 XX Disclosure; Page 2315-2316; 2474pp; English.

XX ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
 CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
 CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, pro hormone activation and neurotransmitter activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing or treating cancers or other
 CC hyperproliferative disorders. Additionally, the secreted proteins and
 CC their nucleic acids may also be used in the treatment of autoimmune
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
 CC wound healing. Nucleic acids of the invention may be used for chromosome
 CC identification, chromosome mapping, in gene therapy, for identifying
 CC individuals from minute biological samples, as hybridisation probes, and
 CC as molecular weight markers. The present sequence represents a human
 CC secreted protein genomic fragment referred to in the disclosure of the
 CC invention

SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match 28.1%; Score 405; DB 7; Length 5033;

Best Local Similarity 100.0%; Pred. No. 6e-146;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTGCGGGCTAGAAAAAATCTGTGCG 60
 DB 13 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTGCGGGCTAGAAAAAATCTGTGCG 72
 QY 61 GTACCAACCCAGAGGGTTGAG 120
 DB 73 GTACCAACCCAGAGGGTTGAG 132
 QY 121 GGACTCAGACTTCCAG 180
 DB 133 GGACTCAGACTTCCAG 192
 QY 181 GCGCTCTCCGCGTCCAG 240
 DB 193 GCGCTCTCCGCGTCCAG 252
 QY 241 ACTGGCGCGCGTGGAG 300
 DB 253 ACTGGCGCGCGTGGAG 312
 QY 301 CGGATGCGCGAG 360
 DB 313 CGGATGCGCGAG 372
 QY 361 GCAGCAAGCTTCCTGCTCATCTTGGCGGGATCCGTCGCACTCG 405
 DB 373 GCAGCAAGCTTCCTGCTCATCTTGGCGGGATCCGTCGCACTCG 417

RESULT 13

AD221005
 ID ADC21005 standard; DNA; 5033 BP.

XX AC ADC21005;

XX 18-DEC-2003 (first entry)

XX Human secreted protein-related DNA sequence #423.

XX gene therapy; human; secreted protein; haemopoietic disorder;
 KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
 KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
 KW leukaemia; wound healing; epithelial cell proliferation disorder;
 KW immune disorder; autoimmune disorder; asthmatic disorder;
 KW cardiovascular disorder; atherosclerosis; myocarditis;
 KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
 KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.

OS Homo sapiens.

XX WO200292787-A2.

XX 21-NOV-2002.

XX 26-MAR-2002; 2002WO-US009257.

XX 27-MAR-2001; 2001US-0278650P.

XX 12-SEP-2001; 2001US-00950082.

XX 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-129287/12.

XX New human secreted proteins and nucleic acid molecules, useful for
 PT preparing a diagnostic or pharmaceutical composition for diagnosing,
 PT preventing or treating hematopoietic or hematologic disorders, e.g.
 PT anemia or hemophilia.

XX PS Disclosure; SEQ ID NO 959; 1512pp; English.

XX CC The invention comprises the amino acid and coding sequences of human

CC secreted proteins. The DNA and protein sequences of the invention are

CC useful for detecting, preventing, diagnosing, prognosticating, treating

CC or ameliorating; haematopoietic or haematological disorders (e.g. anaemia

CC and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease

CC and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);

CC wound healing and disorders of epithelial cell proliferation; immune

CC disorders (e.g. autoimmune disorders and asthmatic disorders);

CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);

CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);

CC and gastrointestinal disorders (e.g. duodenal ulcers and

CC gastroenteritis). The present DNA sequence was used in the

XX exemplification of the invention.

SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match 28.1%; Score 405; DB 9; Length 5033;

Best Local Similarity 100.0%; Pred. No. 6e-146;

Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAAACGGCTACAGACAGTGAGCAAAATAGTTTCGGCTCGCGGCTAGAAAACCTCTGTGC 60

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

13 AAAGTAAACGGCTACAGACAGTGAGCAAAATAGTTTCGGCTCGCGGCTAGAAAACCTCTGTGC 72

QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGGAGAGGTGCA 120

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

73 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGGAGAGGTGCA 132

QY 121 GGAATCAGACTTCACAGCCCACTCGGTCCAGCTTGTAACGAAAGAGACGCCAAGGAC 180

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

133 GGAATCAGACTTCACAGCCCACTCGGTCCAGCTTGTAACGAAAGAGACGCCAAGGAC 192

QY 181 GCGCTCTCCCGGTCAGGCGAGCCAGCTTGCTGGCTTGCTGCTCCCGCTGCGTGCGAGC 240

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

193 GCGCTCTCCCGGTCAGGCGAGCCAGCTTGCTGGCTTGCTGCTCCCGCTGCGTGCGAGC 252

QY 241 ACTCGGCGGCGTGCGAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 300

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

253 ACTCGGCGGCGTGCGAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 312

QY 301 CGGATGCCGCGAGGCTTCAGCGTTCCACTGCTATCGTTATTCAGTGTGTTTGGCTCTA 360

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

313 CGGATGCCGCGAGGCTTCAGCGTTCCACTGCTATCGTTATTCAGTGTGTTTGGCTCTA 372

QY 361 GCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCCACTCG 405

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

373 GCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCCACTCG 417

RESULT 14

ADA53627/c

ID ADA53627 standard; cDNA; 2684 BP.

XX ADA53627;

AC ADA53627;

XX 20-NOV-2003 (first entry)

DT Human coding sequence, SEQ ID 1195.

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;

XX Gene Therapy; human; secretory protein; membrane proteins; cancer;

XX inflammatory disease; osteoporosis; neurological disease; gene; ss.

XX Homo sapiens.

OS EPI293569-A2.

PN 19-MAR-2003.

PD 21-MAR-2002; 2002EP-00006586.

XX PF

XX PS 14-SEP-2001; 2001JP-00328381.

PR 24-JAN-2002; 2002US-0350435P.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-395539/38.

DR P-PSDB; ADA55266.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory

PT and/or membrane proteins, useful for developing medicines for diseases in

PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 1; SEQ ID NO 1195; 205pp; English.

XX The present invention relates to novel human secretory or membrane

CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-

CC ADA54071). The coding sequences are useful in the gene therapy of

CC diseases caused by abnormalities of the proteins, e.g. cancer,

CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 2684 BP; 599 A; 785 C; 684 G; 616 T; 0 U; 0 Other;

SQ Query Match 27.2%; Score 392; DB 7; Length 2684;

Best Local Similarity 100.0%; Pred. No. 6.5e-141;

Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1027 GCGCTCTGTGCTCTTCTCGGAGGCGCGTGGTGGTCTCCAGTATGTTGGCCCCAGC 1086

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

2343 GCGCTCTGTGCTCTTCTCGGAGGCGCGTGGTGGTCTCCAGTATGTTGGCCCCAGC 2284

QY 1087 GCTCTTCGACCCCTTCTGGACCAAGCGCCAGGACTGCGCCAGGAGAGGGGGCTCA 1146

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

2283 GCTCTTCGACCCCTTCTGGACCAAGCGCCAGGACTGCGCCAGGAGAGGGGGCTCA 2224

QY 1147 CCTCTTATCTCGGCGACCCACTGCAAGCAGCGCGCTCTCCAGACTTAAATGTATC 1206

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

2223 CCTCTTATCTCGGCGACCCACTGCAAGCAGCGCGCTCTCCAGACTTAAATGTATC 2164

QY 1207 ACCACTAAGTGGAGGGGACCCCAATCTGGACTCTCTCCCGCCCTTGGGACATCGCAGG 1266

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

2163 ACCACTAAGTGGAGGGGACCCCAATCTGGACTCTCTCTCCCGCCCTTGGGACATCGCAGG 2104

QY 1267 CCGGGAAGCAGTCCCGCCAGGCGCTGGGCGAGGAGCTCCAGGAGGGGACTGAGCGCT 1326

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

2103 CCGGGAAGCAGTCCCGCCAGGCGCTGGGCGAGGAGCTCCAGGAGGGGACTGAGCGCT 2044

QY 1327 GCTGGCGGAGGCGCTCGGACATCCGACGACCCAGGAGGAAAGTCTCTGGGCGGATCTGTA 1386

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

2043 GCTGGCGGAGGCGCTCGGACATCCGACGACCCAGGAGGAAAGTCTCTCTGGGCGGATCTGTA 1984

QY 1387 AATAAACCTTTTCTTTTGTGTTTAAAAA 1418

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1983 AATAAACCTTTTCTTTTGTGTTTAAAAA 1952

RESULT 15

AAS66563/c

ID AAS66563 standard; cDNA; 406 BP.

XX AAS66563;

AC AAS66563;

XX 13-FEB-2002 (first entry)

DT DNA encoding novel human diagnostic protein #2367.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens. 301 CGGCATGCCGCGAGCTTCAGCGTTCCACTGCTCATCGTTATCTAGTGTCTTTGGCTCTA 360
XX OS ||||| 105 CGGCATGCCGCGAGCTTCAGCGTTCCACTGCTCATCGTTATCTAGTGTCTTTGGCTCTA 46
XX PN WO200175067-A2. 361 GCAGCAAGCTTCCTGCTCATCTTGGCGGGGATCCGTTGGCCACTCG 405
XX PD 11-OCT-2001. ||||| 45 GCAGCAAGCTTCCTGCTCATCTTGGCGGGGATCCGTTGGCCACTCG 1
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG02376.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX

Search completed: June 3, 2004, 13:30:00
Job time : 645 secs

Claim 1; SEQ ID NO 2367; 103pp; English.
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activities. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 406 BP; 83 A; 107 C; 138 G; 78 T; 0 U; 0 Other;
Query Match 21.0%; Score 303; DB 5; Length 406;
Best Local Similarity 99.5%; Pred. No. 1.3e-106;
Matches 403; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 AAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGTCGCGGCTAGAAAACTCTGTG 60
Db ||||| 405 AAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGTCGCGGCTAGAAAACTCTGTG 346
61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGGAGGTGCA 120
Db ||||| 345 GGACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGGAGGTGCA 286
121 GGACTCAGACTTCACCAAGCCCACTCGGTCAGCTTGTACGAAAGAGACGCCAGGAC 180
Db ||||| 285 GGACTCAGACTTCACCAAGCCCACTCGGTCAGCTTGTACGAAAGAGACCGTCAAGGAC 226
181 GCGCTCTCCCGGTCCAGGCGAGCCCGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db ||||| 225 GCGCTCTCCCGGTCCAGGCGAGCCCGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 166
241 ACTCGGCGCGGTGCGATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 300
Db ||||| 165 ACTCGGCGCGGTGCGATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 106

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 08:29:22 ; Search time 5809 Seconds
(without alignments)
10751.818 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Result No.	Score	Match	Length	DB	ID	Description
1	1441	100.0	1441	6	AX035346	AX035346 Sequence
2	1029	71.4	1740	9	HSM806241	EX537581 Homo sapi
3	711.4	49.4	1354	10	BC031111	BC031111 Mus muscu
4	626.2	43.5	5851	4	AF547266	AF547266 Sus scrofa
5	465.4	32.3	498	6	AX035348	AX035348 Sequence
6	405.4	28.1	166937	2	AC012255	AC012255 Homo sapi
7	405.4	28.1	181312	9	AC091117	AC091117 Homo sapi
8	403.8	28.0	156534	2	AC009700	AC009700 Homo sapi
9	403.8	28.0	171444	2	AC087790	AC087790 Homo sapi
10	396	27.5	2684	6	AX714511	AX714511 Sequence
11	396	27.5	2684	9	AK056896	AK056896 Homo sapi
12	396	27.5	156534	2	AC009700	AC009700 Homo sapi
13	356.4	24.7	1636	6	AX405919	AX405919 Sequence
14	356.4	24.7	1923	9	BC029819	BC029819 Homo sapi
15	348.8	24.2	1521	10	BC019755	BC019755 Mus muscu
16	299.4	20.8	2439	9	AK027187	AK027187 Homo sapi
17	245.8	17.1	1706	9	BC020841	BC020841 Homo sapi
18	245.8	17.1	2684	6	AX714511	AX714511 Sequence
19	245.8	17.1	2684	9	AK056896	AK056896 Homo sapi
20	217	15.1	217	9	HS183H12F	Z57552 H.sapiens C
21	199.8	13.9	218	9	HS183A12R	Z59354 H.sapiens C
22	167.8	11.6	229583	10	AL844566	AL844566 Mouse DNA
23	166	11.5	1923	9	BC029819	BC029819 Homo sapi
24	165.8	11.5	231868	2	AC118124	AC118124 Rattus no
25	159.4	11.1	166937	2	AC012255	AC012255 Homo sapi
26	159.4	11.1	171444	2	AC087790	AC087790 Homo sapi
27	159.4	11.1	181312	9	AC091117	AC091117 Homo sapi
28	154.2	10.7	231868	2	AC118124	AC118124 Rattus no
29	147	10.2	229583	10	AL844566	AL844566 Mouse DNA
30	123.2	8.5	110250	2	AC138220	AC138220 Mus muscu
31	121.6	8.4	55061	2	AC091597	AC091597 Mus muscu
32	101	7.0	110250	2	AC138220	AC138220 Mus muscu
33	65.6	4.6	2207	3	AY061346	AY061346 Drosophil
34	63.4	4.4	125020	9	AF429315	AF429315 Homo sapi
35	61	4.2	7218	6	I66494	I66494 Sequence 14
36	60.6	4.2	2000	6	AX655393	AX655393 Sequence
37	59.6	4.1	125020	9	AF429315	AF429315 Homo sapi
38	56.8	3.9	63689	2	AC024268	AC024268 Homo sapi
39	55.8	3.9	4615	1	STWHIB12X	L22864 Streptomyce
40	55.2	3.8	151173	8	AC135257	AC135257 Genomic s
41	53	3.7	2690	8	AY332473	AY332473 Oryza sat
42	53	3.7	169162	8	AP004267	AP004267 Oryza sat
43	52	3.6	2000	6	AX655393	AX655393 Sequence
44	52	3.6	117787	2	AC104271	AC104271 Oryza sat
45	51.8	3.6	168866	8	OSJN00023	AL606588 Oryza sat

ALIGNMENTS

RESULT 1	AX035346	AX035346	1441 bp	DNA	linear	PAT 15-NOV-2000
LOCUS	Sequence 1 from Patent WO0053748.					
DEFINITION	AX035346					
ACCESSION	AX035346					
VERSION	AX035346.1	GI:11191064				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.					
TITLE	Novel compounds					
JOURNAL	Patent: WO 0053748-A 1 14-SEP-2000;					

Pred. No. is the number of results predicted by chance to have a

BRUCK CLAUDE ELVIRE MARIE (BE) ; SMITHLINE BEECHAM BIOLOG (BE) ;
VINALS Y DE BASSOLS CARLOTA (BE) ; COCHE THIERRY (BE) ; CASSART
JEAN POL (BE)

FEATURES

Location/Qualifiers
1..1441
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1441; DB 6; Length 1441;
Best Local Similarity 100.0%; Pred. No. 3.3e-294;
Matches 1441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTG 60
1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTG 60
61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGGAGAGGTGCA 120
61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGGAGAGGTGCA 120
121 GGAATCAGACTTACAGCCAGCTCGGTCCAGCCCTTGTACGCAAGAGAGCCCAAGGAC 180
121 GGAATCAGACTTACAGCCAGCTCGGTCCAGCCCTTGTACGCAAGAGAGCCCAAGGAC 180
181 GCGCTCTCCCGGTCAGGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
181 GCGCTCTCCCGGTCAGGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
241 ACTCGGCGGCTGAGCAGTCAACCTGTTGGAACGGGCTACTGCTTTTACCCCGAGCCC 300
241 ACTCGGCGGCTGAGCAGTCAACCTGTTGGAACGGGCTACTGCTTTTACCCCGAGCCC 300
301 CGGATGCGCGAGGCTTCAAGCTTCCACTGCTCATGCTTATTTAGTGTGCTGCTCA 360
301 CGGATGCGCGAGGCTTCAAGCTTCCACTGCTCATGCTTATTTAGTGTGCTGCTCA 360
361 GCAGCAAGCTTCTGCTCATGCTTCCGGGATCCGTCGCTGCTGCTGCTGCTGCTGCTG 420
361 GCAGCAAGCTTCTGCTCATGCTTCCGGGATCCGTCGCTGCTGCTGCTGCTGCTGCTG 420
421 GTGAGAGTTCTTCTCAGTCTGTTTATAGGCGAGAAATTTGCTGCTGCTGCTGCTGCTG 480
421 GTGAGAGTTCTTCTCAGTCTGTTTATAGGCGAGAAATTTGCTGCTGCTGCTGCTGCTG 480
481 GAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
481 GAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
541 GTTACAGCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
541 GTTACAGCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
601 ACCCAGTGCATCAGCTGAAACGAGACCATGACTCAACAGGAGTTCACTGCGCTCTG 660
601 ACCCAGTGCATCAGCTGAAACGAGACCATGACTCAACAGGAGTTCACTGCGCTCTG 660
661 AAGAGAAATTACGCGCGGAGTACCGAAACGCTGAGAGAGGGGCTGCGGACCCAGTG 720
661 AAGAGAAATTACGCGCGGAGTACCGAAACGCTGAGAGAGGGGCTGCGGACCCAGTG 720
721 CTCTACCTGGCGGAGAGTTTCAACGAGTAGCCCTTGGGCTGCTGCTGCTGCTGCTGCTG 780
721 CTCTACCTGGCGGAGAGTTTCAACGAGTAGCCCTTGGGCTGCTGCTGCTGCTGCTGCTG 780
781 CTGGCGGAGACATACGCTCGGCGAGCTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
781 CTGGCGGAGACATACGCTCGGCGAGCTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
841 AACGTGCTGCTCTCCAGCGCGGCTCTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
841 AACGTGCTGCTCTCCAGCGCGGCTCTACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

901 GCCTTCGGCTCTTTCGGGCTCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 960
901 GCCTTCGGCTCTTTCGGGCTCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 960
961 CTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 1020
961 CTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 1020
1021 GCAACCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1080
1021 GCAACCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1080
1081 CCCAGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 1140
1081 CCCAGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 1140
1141 GGCTCACCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 1200
1141 GGCTCACCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 1200
1201 TGTATCACCACTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1260
1201 TGTATCACCACTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1260
1261 CGCAGCGCGGAGAGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 1320
1261 CGCAGCGCGGAGAGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 1320
1321 AGCGTGTCTGCGGAGAGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1380
1321 AGCGTGTCTGCGGAGAGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1380
1381 TCTGTAATAAATTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1440
1381 TCTGTAATAAATTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1440
1441 A 1441
1441 A 1441

RESULT 2

HSM806241
LOCUS HSM806241 1740 bp mRNA linear PRI 17-JUN-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686C04213 (from clone DKFZp686C04213).
ACCESSION BX537581
VERSION BX537581.1 GI:31873659
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1740)
AUTHORS Bloembergen, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686C04213) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/
location/Qualifiers
1..1740
/organism="Homo sapiens"
/mol_type="mRNA"

```

/db_xref="taxon:9606"
/map="15q15"
/clone="DfXZp686C04213"
/tissue_type="human rectum tumor"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
polyA_signal 1694.1699
polyA_site 1720
ORIGIN

```

ORIGIN

Query Match	71.4%;	Score 1029;	DB 9;	Length 1740;
Best Local Similarity	98.6%;	Pred. No. 4.5e-207;		
Matches 1038; Conservative		0; Mismatches 15;	Indels 0;	Gaps 0;

Qy	381	CTTCCCGGGATCCGTGGCCACTCGCGCTGGTTTGGTTGGTGAGAGTTCTTCTCAGTCT	440
Dd	698	CTGACCCGGGTGCCATTCTCTGCAGCGCTGGTTTGGTTGGTGAGAGTTCTTCTCAGTCT	747
Qy	441	GTTCTAGCGCCGAGAAATTGTGGCTGTGCACCTTCAGTGCAGAAATGGTTCTGGGTACAGT	500
Dd	748	GTTCTAGCGCCGAGAAATTGTGGCTGTGCACCTTCAGTGCAGAAATGGTTCTGGGTACAGT	807
Qy	501	GAACACCAACACATCCTTACAAGCCCTTCAGCGCAGCGCGCTTACAGCCCGGTGCGGTCT	560
Dd	808	GAACACCAACACATCCTTACAAGCCCTTCAGCGCAGCGCGCTTACAGCCCGGTGCGGTCT	867
Qy	561	GCTCGTGGCCCTGGAGGSCATTAAATATTACACTCACAGGACCCAGTGCATCAGCTGAA	620
Dd	868	GCTCGTGGCCCTGGAGGSCATTAAATATTACACTCACAGGACCCAGTGCATCAGCTGAA	927
Qy	621	CGAGACCATTGACTACBACGACAGTTCACTGGCGTCTGAAAGAGAATTACGCCGCGGA	680
Dd	928	CGAGACCATTGACTACBACGACAGTTCACTGGCGTCTGAAAGAGAATTACGCCGCGGA	987
Qy	681	GTACCGGAACGCACCTGGAGAAGGGCTGCCGGAACCACTGTCTTACCTGGCGGAGAAGTT	740
Dd	988	GTACCGGAACGCACCTGGAGAAGGGCTGCCGGAACCACTGTCTTACCTGGCGGAGAAGTT	1047
Qy	741	CACACCGAGTAGCCCTTGCGCCCTGTACCAACAGTACCACTGGCGGGAACACTACGCTTC	800
Dd	1048	CACACCGAGTAGCCCTTGCGCCCTGTACCAACAGTACCACTGGCGGGAACACTACGCTTC	1107
Qy	801	GGCCACGCTATGGGTGGCGTTCTGCTTCTGGCTCCTCTCCAAACGTGTGCTCTCCAGGCC	860
Dd	1108	GGCCACGCTATGGGTGGCGTTCTGCTTCTGGCTCCTCTCCAAACGTGTGCTCTCCAGGCC	1167
Qy	861	GGCCCCGCTCTACGGAGGCCCTGGCACCTGTGACCAACGGAGCCCTTCGCGCTCTTCGGGGT	920
Dd	1168	GGCCCCGCTCTACGGAGGCCCTGGCACCTGTGACCAACGGAGCCCTTCGCGCTCTTCGGGGT	1227
Qy	921	CTTTCGCTTGGCTCCATCTCTAGCGTGCCGCTCTGCCCCGTCCGCCCTAGGCTCTTCGCG	980
Dd	1228	CTTTCGCTTGGCTCCATCTCTAGCGTGCCGCTCTGCCCCGTCCGCCCTAGGCTCTTCGCG	1287
Qy	981	GCTCACCACTCAGTACGGCGCCCTTCTGGGTCAAGTGGCAACCGGGCTCCTGTGCT	1040
Dd	1288	GCTCACCACTCAGTACGGCGCCCTTCTGGGTCAAGTGGCAACCGGGCTCCTGTGCT	1347
Qy	1041	CTTTCCTCGGAGGGCCGTGGTGAGTCTCCAGTATGTTTCGGCCACCGCTCTTCGCACCTT	1100
Dd	1348	CTTTCCTCGGAGGGCCGTGGTGAGTCTCCAGTATGTTTCGGCCACCGCTCTTCGCACCTT	1407
Qy	1101	TCTGGACCAAGCGCCAAAGCATGACAGCCAGGAGAGGGGGCTCACCTCTTATCCTCGG	1160
Dd	1408	TCTGGACCAAGCGCCAAAGCATGACAGCCAGGAGAGGGGGCTCACCTCTTATCCTCGG	1467
Qy	1161	CGACCCACTGCACAAGCAGGCCGCTCTCCAGACTTAAATGTATCACCACTAACCTGTG	1220
Dd	1468	CGACCCACTGCACAAGCAGGCCGCTCTCCAGACTTAAATGTATCACCACTAACCTGTG	1527
Qy	1221	AGGGGACCCAAATCTGGACTCTCTTCCCGCTTCGGGACATCGCAGGCCCGGGAAGCAGTGC	1280

Db	1528	AGGGGACCCAACTGGA	CTCCTTCCCCCGCTTGGGACATCGCAGCGCCGGGAAGCAGTGC	1587
Qy	1281	CCGCCAGGCTGGGCCAGGAGAGCTCCAGGAAGGCCACTGAGCGCTGCTGGCGGAGGCC	1340	
Db	1588	CCGCCAGGCTGGGCCAGGAGAGCTCCAGGAAGGCCACTGAGCGCTGCTGGCGGAGGCC	1647	
Qy	1341	TCGGACATCCGACGGCACCCAGGGAAGTCTCTCTGGGGGCGATCTGTAAATAAACCTTTTT	1400	
Db	1648	TCGGACATCCGACGGCACCCAGGGAAGTCTCTCTGGGGGCGATCTGTAAATAAACCTTTTT	1707	
Qy	1401	TCCTTTTGTTTTTTAAAAA	AAAAA 1433	
Db	1708	TCCTTTTGTTTTTTAAAAA	AAAAA 1740	

RESULT 3	DEFINITION
BC031111	
LOCUS	

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE AUTHORS

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

BC031111 1354 bp mRNA linear ROD 04-NOV-2003
Mus musculus RIKEN CDNA 9030623N16 gene, mRNA (CDNA clone MGC:35715
IMAGE:4989815), complete cds.
BC031111
BC031111.1 GI:21411407
MGC.

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 {bases 1 to 1354}

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.F., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalusz, D.E.,
Schein, J., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 157,000
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2238257
12477932
2 (bases 1 to 1354)
Strausberg, R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
cDNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>

Series: IRAK Plate: 58 Row: g Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13385245.

FEATURES

Location/Qualifiers
1..1354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:35715 IMAGE:4989815"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_lib="NCI CGAP_Co24"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"

gene

1..1354
/gene="9030623N16Rik"
/db_xref="LocusID:66811"
/db_xref="MGI:1914061"

CDS

129..1091
/codon_start=1
/product="RIKEN cDNA 9030623N16"
/protein_id="AAH31111.1"
/db_xref="GI:21411408"
/db_xref="LocusID:66811"
/translation="MTAWDGVLPFPQPRHAASFVPLLVILVFLSLAASFILPLP
IRGSRWFLVRLVLSLPIGAEIVAVHFSGDWFGVWNTSVKAFSPSRVQVHVGLH
VGLAGVNTLRGTPLQNLNETIDYNERFTWRLNEDYTKYVHALEKGLPDPVLYAEK
FTPSSPCGLYHQVHLGAYAAATLVAFCEWIIANALLSMPAPLYGGLALLTTGAFTL
FGVAFASISSVPLCHPRLGSAVLTPYVYASFWLTATGILSLLLGGAVVILHYTRPS
AIRFLDLSVKDCSNQAKGNSPLTNNPQHEQLKSPDLNITLL"

ORIGIN

Query Match 49.4%; Score 711.4; DB 10; Length 1354;
Best Local Similarity 75.9%; Pred. No. 6.6e-140;
Matches 936; Conservative 0; Mismatches 281; Indels 16; Gaps 4;

QY	222	CTGCCCGCCTCGTGCAGCACTCGCCCGCGTGCAGCATGACCTGTGGACGGGTACT	281
Db	92	CTCGCTGCCTCGCAACAGTTCGNACTTCTTCAGAGATGACTGCGTGGGACGGGTGCT	151
QY	282	GCCTTTTACCCCGACCCCGGATGCCCGAGCTTCAGCGTTCAGCTGCTCATCGTTAT	341
Db	152	ACCCCTTACCCCGACCCCGGATGCCCGAGCTTCAGCGTTCAGCTGCTCATCGTTAT	211
QY	342	TCTAGTGTGTTTGGTCTAGCAGCAGCTTCCTGCTCATCTTCCCGGGGATCCCGGCCA	401
Db	212	CCTGCTGTTCTTGTCTTGGTCTAGCTTCCTGTTTCATCTTGCCTGGAATCCCGTGGCA	271
QY	402	CTCGCGCTGTTTGGTGTGAGAGTTCCTCTCAGTCTGTTTCATAGGCGCAGAAATGT	461
Db	272	CTCGCGCTGTTTGGTGTGAGAGTTCCTCTCAGTCTGTTTCATAGGCGCAGAAATGT	331
QY	462	GGCTGTGCACCTTCAGTGCAGAAATGGTTCGTGGGTACAGTGAACACCAACACATCTACAA	521
Db	332	GGCTGTGCACCTTCAGTGCAGAAATGGTTCGTGGGTACAGTGAACACCAACACATCTACAA	391
QY	522	AGCCTTCAGCGCAGCGCGGTACAGCCCGTGTGCGTCTGCTGCGGCGCTGGAGGGCAT	581
Db	392	AGCCTTCAGTCCATCCCGTGTTCAGTCCATGTCGCTGTGCGGCGCTGGAGGGCGT	451
QY	582	TAATATTACACTCACAGGACCCCGTGCATCAGCTGAACGAGACCAATTGACTACACGA	641
Db	452	TAACATTACACTCCGAGGAACACCGAGGACAGTGAACGAGACCAATTGACTACCAATGA	511
QY	642	GCAGTTCACTGGCGTCTGAAGAAGAAATPACGCCCGGAGTACGCGAACCGCACTGGAGAA	701
Db	512	GCCTTCACTGGCGTCTGAACGAAGACTACACCAAGGAGTATGTCATGCTCCCTGGAGAA	571
QY	702	GGGCTGCCGAGCCCGTGTCTACCTGGCGGAGAGTTACACCGAGTAGCCCTTGGCG	761
Db	572	GGGACTGCCGAGCCCGTGTCTACCTGGCAGAGAGTTACACCGAGTAGCCCTTGGCG	631
QY	762	CCTGTACCAACAGTACCACTGGCGGGGACACTAGCCCTCGGCCACGCTATGGTGGCGTT	821

Db	632	GTGTACCAACCAATATCACCTGCGGTCACTATGCGCGCAACACTGTGGTGGCATT	691
QY	822	CTGTTCTGGCTCTCTCCAAAGTGTGCTCTCCACGCGCGCGCTCTACGAGGCGCT	881
Db	692	CTGCTTCTGATCATCGCAATGCGTGTCTCTCATGCGCGCCACTCTACGAGGCGCT	751
QY	882	GGCACTGTGACCAACCGAGCGCTTCGCGCTCTTCGGGGTCTTCGCTTGGCGCTCATCTC	941
Db	752	GGCTTGTGCTCACCAACCGGTGCTTCACGCTCTTCGGTGTCTTCGCTTGGCTTCGATTC	811
QY	942	TAGCGTCCCGCTCTGCGCTCGCTCGCTAGGCTCTTCGCGCTCACCACTCAGTACGCGC	1001
Db	812	CAGCGTCCCGCTCTGCGCACTTCGCGCTGGGCTTCGCGCTCTCACGCTTACTACGCGC	871
QY	1002	CGCTTCTGGGTCACTGCGTGGCAACCGGCGTCTCTGCTGCTCTCTCGGAGGGCGGTGT	1061
Db	872	CTCCTTGTGGTCACTGCGTGGCAACCGGCTCTCTGCTGCTCTCTCGGAGGGCGGTGT	931
QY	1062	GAGTCTCCAGTATGTTGCGCGCCAGCGCTCTTCGCAACCTCTTCGACCAAGCGCCAGCA	1121
Db	932	GATTCTCCACTATCTGCGCGCCAGCGCTCTGCTGCTCTCTGATCTAAGTGTCAAAGA	991
QY	1122	CT---CGAGCCAGGAGAGGGGCTCACCTCTTATCTCTGCGGACCCACTGCACAGCA	1178
Db	992	CTGTAGCAACAGGCTAAAGGAACTCACCTCTCACCTCAACAACCCGCAACACGAGCA	1051
QY	1179	GGCGCTCTCCAGACTTAAATGTATCACCACTAACCTGTGAGGGGACCCCAATCTGGA	1238
Db	1052	GTTGAGAGTCCAGACTTAA---TATTACCACTCTCTCTGTGAAATAAATGACTCCGGA	1108
QY	1239	CTCCTTCCCGCTTGGGACATCGCAGCGCGGGAAGCAGTCCCCCGCAGGCGCTG-GGCCA	1297
Db	1109	TTTCTACCCCTCTTTGGGACCCCATAGACCTGAAGACTGTGTTAAGAGCGGTGCCAGGA	1168
QY	1298	GGAGAGTCCAGGAAGGCACTGAGCGCTGTGCGGCGAGGCGCTTCGACATCCCGAGGCA	1357
Db	1169	GCATAGCTGAGTCTGGAGAACTGCTTCTCTCCCTGCGGAGCAGCAAGGGGGACCCCATGTG	1228
QY	1358	CC-----AGGAAAGTCTCTGCGGCGCATCTGTAAATAAAACCTTTTCTTTTGT	1408
Db	1229	CTTTTGTCTAGAGAAATGATTTCTCAGGAGAACTGTAAATAAACTTTTGTCTTTT	1288
QY	1409	TTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1441
Db	1289	TTTTTTCTTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1321

RESULT 4

AF547266
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

AF547266
Sus scrofa dual oxidase 1 (Duox1) mRNA, complete cds.
AF547266.1
GI:23664372
Sus scrofa (pig)
Sus scrofa
Sus scrofa
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 5851)
Morand, S., Chaaroui, M., Kaniewski, J., Deme, D., Ohayon, R.,
Noel-Hudson, M.S., Virion, A. and Dupuy, C.
Effect of iodide on nicotinamide adenine dinucleotide phosphate
oxidase activity and Duox2 protein expression in isolated porcine
thyroid follicles
Endocrinology 144 (4), 1241-1248 (2003)
22526370
12639906
2 (bases 1 to 5851)
Kaniewski, J., Morand, S., Noel-Hudson, M.-S., Ohayon, R., Virion, A.
and Dupuy, C.
Direct Submission
Submitted (13-SEP-2002) Unite 486, INSERM University of Paris-Sud,

5 rue JB Clement, Chatenay-Malabry 92296, France

Location/Qualifiers

1.5851

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

1.5851

/gene="Duox1"

/gene="Duox1"

948.5609

/EC_number="1.6.3"

/function="putatively generates H2O2 by transferring electrons from NADH/NADPH to molecular oxygen"

/function="oxidoreductase acting on NAD(P)H with oxygen as acceptor"

/note="NADPH:O2 oxidoreductase"

/codon_start=1

/product="dual oxidase 1"

/protein_id="AAN39338.1"

/db_xref="GI:23664373"

/translation="MGFRLALAWTLTVGPWPMGMGARNSTISWEVQRFDGWNNLMHEKHG
GSKGSRQLRVPASVADGVQPLGEPHLPNPRDLSTNMRGSPAGQASLRNRTVLGVFF
GYHVLSDLVSTIEKPGCPAEFLNIHIPGDPVDFPHKSGDVVLPPQRSRWDPNTGQSPS
NPRDLTNEVTGWLGDGAJYGSSSHWSDELSFSGQLASGDPAPFPAQADPLFWMTLP
PDPATQGRPGQIYAFGAEGQNRPEFLQALGLLWFRFYHNLCAQKLAREHPLWGDEELP
QHARKVIATQSIITMYEWLPSFLRKMPEYTYGYPFLDPSISPEFLAAEQFFSTMV
PPGVYMRNASCHFGQVINRNSVSRALRVCSYWSREHPNMQRAEDVDALLGEMASQI
AEREDHMVEDVDQFWPGPLKPSRTDHLASCLQRDLGLPFSYTKARARGLGPPVTRV
QDINPALSRSQIIVLEATAALYNQDLSRLLELPGELLESYGDGPGPLFSTIVLDQFVRL
RDGDRYWFTENKGLFSEKETAEIRNTSLRDVIAVNTMTGALQPNVFFWHAQDPCP
QPROLSTKDLPAACAPLIMRDYFKSGGFGVYTGTLCCFPLVSLLSAWIYAQLRRRNF
KRLPQQVNRSLTMEKLVGGMKALEWQKEPCRPVLVHLQSGQIHVMDGRLSVLRTIQ
ILRPQQVNRSLTSSNHGRTLLIKIPKEYDLVLMFDEEERQVMVENLQSAKESGLSF
QEWELREQLMRAAVTREQRSHLLETFFRHLSQVLDIDQADAGALPLDSSQKVREAL
TCELSRAEEAESIIGLKQDMFVBSMFSLADKDGNGYLSFREFLDILVFMKGSPEKS
RLMFRMYDFDGNLISKDEFIRMLRSFIEISNNCLSKAQLETVVESMFREACFQDQKE
LTWEDFHFLMRDSDSELFTQLCVKGVVEVPIKDLCRRASTYIQEKLCPSPRVSAHK
PRSNVDVEVELTPWKLCQPTDTPPQEIIRRRFCKKVTSFQPLIFTFAHREKFORSRH
QTWQPKRFVENYRRHIGCLAVPYTIAGGLFLERAYYAFAAHMGITDTRVGIIIS
RGTAASISFMFSYILLTMCRNLIITFLRETFLNRVVPFDDAAVDFHRLIASTAILITVLH
SAGHVNVLVFSISPLSVSLCLPGLFDHNGSEFPQKYIYWFPTQVPLGTGWMLLLI
AIMYVFASSHFRCSFRGFWLTHELYILLYMLLIHGSFALIQLPFHIFFLVPALTIY
VGDKLVSLSRCKVEISVWKAELLPSGVTHLQFORPQGSFYKSGQWRIACLAGTTEY
HPTLTSA PHEDTILSLHRAAGPWTTRLREIYSPPTDNCAYPKLYLDGPFGEHQH
WHKFEVSLVGGIGVTPPASILKDLVFKSSVSCQVCKKIYIIVWTRTQRQFENLAD
IIREVBENDHRDLVSVHIYITQLAEKFDLRTMLYICERHFKQVLRSLFTGLRSITH
FGRRPFPFFNSLQEVHPQVRKIVGFSCGPPGTMKNVEKACQLINRQDRTFESHYEN
P"

ORIGIN	Query Match	43.5%;	Score 626.2;	DB 4;	Length 5851;
	Best Local Similarity	81.8%;	Pred. No. 7.6e-122;		
	Matches 735;	Conservative	0;	Mismatches 163;	Indels 1; Gaps 1;
QY	254	GCAGCATGACCTGTGTGAAACGGCGTACTGCGCTTTTACCCCGACGCCCGGCATGCCGCAG	313		
DB	10	GCAGCATGACTCTGTGTGAAACGGCGTGCTGCCCTTCTACCCCTAGCCCTCGGTATGCCGCCA	69		
QY	314	GCTTCAGCGTTCCACTGCTCATCGTTATTCCTAGTGTCTTTTGGCTCTAGACGAAGCTTCC	373		
DB	70	GCTTCAGCGTCCCACTACTCATCGTTATTCCTGGTGTCTTTGGCCTTAGCCACCAGCTTTC	129		
QY	374	TGCTCATCTTGGCCGGGGATCCGTGGCCCACTCGGGCTGTTTTGGTTGGTGAGAGTTCCTC	433		
DB	130	TACTCATCTTGGCCGGGATTCGTGGCCCACTCGGGCTGGTTCTGGTTGGTGAGAGTTCCTC	189		
QY	434	TCAGTCTGTTCATAGCGCGCAGAAATTGTGGCTGTGCACCTTCAGTGCAGAAATGGTTCGTGG	493		
DB	190	TCAGCCTGTTCTATAGGGCGCAGAAATTGTGGGGTGTCACTTCAGCGCAGAAATGACCCGTGG	249		
QY	494	GTACAGTGAACACCAACACATCCTACAAAGCCCTTCAGCGCAGCGCGCTTACAGCCCCGTG	553		
DB	250	GCGAAGTTAGCAATAACATCTTACAAGGCCCTTCAGTGTGGCAGCTGTCCAGCCCCACC	309		

Qy	554	TGGTCTGCTCGTGCCCTGGAGGCATTAAATTTACACTCACAAGGACCCCAGTCATC	613
Dd	310	TGGTCTGCACTGGGCCCTGGAGGTGTGAATATCACACTCACAAGGAATCCGGTACAGC	369
Qy	614	AGCTGAACGAGACCATTGACTACAACGAGCAGTTACCTGGCGTCTGAAGAGAATAACG	673
Dd	370	AGCTGAATGAGACCATCGACTACAACGAGCAGTTATCTGGCGTATGGCGAAAATATG	429
Qy	674	CCGCGAGTAGCGGAACGACCTGGAGAACGGGCTGCCGACCCAGTCTCTACTCTGGCGG	733
Dd	430	CGCGCGGTATGCGAAGGCTCTGGAGAACGGGCTGCCGACCCAGTCTCTACTCTGGCGG	489
Qy	734	AGAAGTTACACACCGAGTAGCCCTTTCGGCTCTGTAACCAACGATACCACTGGCGGACACT	793
Dd	490	AGAAGTT-ACTCCAGAACCCCCCTGCCGGGTTTACCGCCAATATGCGCTAGCCGACACT	548
Qy	794	ACGCTCGGCCACCGTATGGGTGGGTTCTGCTTCTGGCTCTCTCAACGTCGCTCT	853
Dd	549	ACGCTCGGCCACATGTGGAAOOGGCTTCTGCTTCTGACTCTCTCAACCTGCTGCTTT	608
Qy	854	CCAGCGGCGCCGCTCTACGGAGCGCTGGCACTGCTGACCAACGAGACCTTCGCGCTCT	913
Dd	609	CCATGCGGTCGCCGATATGGAGGCTGGCTCTCTCTCATCACCGCGGCTTCGCGCTCT	668
Qy	914	TGGGGTCTTGGCTTGGCTCCATCTCTAGCGTGGCTCTGCGCTCTGCGCTCTAGGCT	973
Dd	669	TCTCTGTCTTGGCTTGGCTCTATCTCCAGGCTGCTCTTTGCCAGCTCCGCTCGGCT	728
Qy	974	CCTCCGCGCTCACCACTCAGTACGGCGCGCTTCTGGGTTCAGCTGSCAACCGCGCTCC	1033
Dd	729	CCTCCGAGCTCACCACTCAGTACGGCTGCTGCTCTTTGGATCACGCTGGCCACCGGCTCC	788
Qy	1034	TGTGCTCTTCTCGGAGGCGCGTGGTGGTCTCAGTATGTTGGCGCCAGCGCTCTTC	1093
Dd	789	TGTGCTCTTCTCGGAGGCGCGTGGTGGTCTCAGTATGTTGGCGCCAGCGCTCTTC	848
Qy	1094	GCACCTTCTGGACCAAAGCGCAAGACTGCAGCCAGGAGAGGGGGCTCACCTCTT	1152
Dd	849	GCACCTTCTTAGAGGAGATATCAAGGACTGCGATAGTCAGGCAAGGGTCTCCATTT	907

RESULT 5
AX035348
LOCUS AX035348 498 bp DNA linear PAT 15-NOV-2000

DEFINITION Sequence 3 from Patent WO0053748.

ACCESSION AX035348

VERSION AX035348.1 GI:11191065

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.
Novel compounds
Patent: WO 0053748-A 3 14-SEP-2000;

TITLE BRUCK CLAUDE ELVIRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ;
JOURNAL VINALS Y DE BASSOLS CARLOTA (BE) ; COCHE THIERRY (BE) ; CASSART JEAN POL (BE)

FEATURES Location/Qualifiers
source 1..498
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 32.3%; Score 465.4; DB 6; Length 498;
Best local Similarity 99.4%; Pred. No. 6.6e-88;
Matches 488; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 939 CTCTAGCGTGGCGCTCTGCGGCT-CGCGCTAGGCTCTCGGCTCACCACTCAGTACG 997

ORIGIN

Query Match	32.3%	Score 465.4;	DB 6;	Length 498;
Best local similarity	99.4%	Pred. No. 6.e-88;		
Matches 488; Conservative	0;	Mismatches 1;	Indels 2;	Gaps 2;
Qy	939	CTCTAGCGTGC CGCTCTG CCGCT-CCGCGCTAGGCTCTCTCCGCGCTCACCACTCAGTACG	997	

QY 939 CTCTAGCGTCCGGCTCTGCGGCT - CCGCTAGGCTCTCCGGCTCACCACTAGTAGC 997

Db 1 CTCTAGCGTCCGCTCTGCGCGCTCCCGCTAGGCTCTCTCGCGCTCACCACCTCAGTACG 60

QY 998 -GCGCGCCCTTCTGGGTACGCTGGCAACCGCGTCTCTGTCCTCTTCTCGGAGGGCC 1056

Db 61 AGCGCGCCCTTCTGGGTACGCTGGCAACCGCGTCTCTGTCCTCTTCTCGGAGGGCC 120

QY 1057 GTGGTGAGTCTCCAGTATGTTGGGCCCGGCTCTTTCGACCCCTTCTGGACCAAGCGCC 1116

Db 121 GTGGTGAGTCTCCAGTATGTTGGGCCCGGCTCTTTCGACCCCTTCTGGACCAAGCGCC 180

QY 1117 AAGGACTGCAGCCAGGAGAGAGGGGGCTCCTCTTATCTCGGCGACCCACTGCACAAG 1176

Db 181 AAGGACTGCAGCCAGGAGAGAGGGGGCTCCTCTTATCTCGGCGACCCACTGCACAAG 240

QY 1177 CAGGCGCGCTCTCCAGACTTAAATATATCACTAACTCTGAGGGGGGACCCCAATCTG 1236

Db 241 CAGGCGCGCTCTCCAGACTTAAATATATCACTAACTCTGAGGGGGGACCCCAATCTG 300

QY 1237 GACTCCTTCCCGCCTTGGGACATCGCAGGCGGGAAGCAGTGCCTCCCGCCTGGGCC 1296

Db 301 GACTCCTTCCCGCCTTGGGACATCGCAGGCGGGAAGCAGTGCCTCCCGCCTGGGCC 360

QY 1297 AGGAGAGCTCCAGGAGGGCACTGAGCGCTGCTGGCGGAGGCGCTCGGACATCGCAGGC 1356

Db 361 AGGAGAGCTCCAGGAGGGCACTGAGCGCTGCTGGCGGAGGCGCTCGGACATCGCAGGC 420

QY 1357 ACAGGGGAAAGTCTCTGGGGGAGTCTGTAAATAAACCTTTTCTTTTGTTTTAAA 1416

Db 421 ACAGGGGAAAGTCTCTGGGGGAGTCTGTAAATAAACCTTTTCTTTTGTTTTAAA 480

QY 1417 AAAAAAATAAAA 1427

Db 481 AAAAAAATAAAA 491

RESULT 6

AC012255

LOCUS

DEFINITION Homo sapiens chromosome 15 clone RP11-109D20 map 15, WORKING DRAFT

SEQUENCE, 22 unordered pieces.

AC012255

AC012255.4 GI:7684440

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 3, 2000 this sequence version replaced gi:7209928.

All repeats were identified using RepeatMasker:

Smit, A.P.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3737

Center clone name: 109 D 20

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 153800 bases at least Q40

Consensus quality: 160606 bases at least Q30

Consensus quality: 163275 bases at least Q20

Insert size: 164837; sum-of-contigs

Quality coverage: 4.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 22 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 2087: contig of 2087 bp in length

* 2088 2187: gap of 100 bp

* 2188 3626: contig of 1439 bp in length

* 3627 3726: gap of 100 bp

* 3727 4845: contig of 1119 bp in length

* 4846 4945: gap of 100 bp

* 4946 6702: contig of 1757 bp in length

* 6703 6802: gap of 100 bp

* 6803 8326: contig of 1524 bp in length

* 8327 8426: gap of 100 bp

* 8427 10494: contig of 2068 bp in length

* 10495 10594: gap of 100 bp

* 10595 13215: contig of 2621 bp in length

* 13216 13315: gap of 100 bp

* 13316 16451: contig of 3136 bp in length

* 16452 16551: gap of 100 bp

* 16552 20843: contig of 4292 bp in length

* 20844 20943: gap of 100 bp

* 20944 24940: contig of 3997 bp in length

* 24941 25040: gap of 100 bp

* 25041 29901: contig of 4861 bp in length

* 29902 30001: gap of 100 bp

* 30002 34189: contig of 4188 bp in length

* 34190 34289: gap of 100 bp

* 34290 38386: contig of 4097 bp in length

* 38387 38486: gap of 100 bp

* 38487 45827: contig of 7341 bp in length

* 45828 45927: gap of 100 bp

* 45928 51729: contig of 5802 bp in length

* 51730 51829: gap of 100 bp

* 51830 60123: contig of 8294 bp in length

* 60124 60223: gap of 100 bp

* 60224 73532: contig of 13309 bp in length

* 73533 73632: gap of 100 bp

* 73633 87433: contig of 13801 bp in length

* 87434 87533: gap of 100 bp

* 87534 101708: contig of 14175 bp in length

* 101709 101808: gap of 100 bp

* 101809 121193: contig of 19385 bp in length

* 121194 121293: gap of 100 bp

* 121294 139051: contig of 17758 bp in length

* 139052 139151: gap of 100 bp

* 139152 166937: contig of 27786 bp in length.

FEATURES

Source

1..166937

/organism="Homo sapiens"

/mol_type="genomic DNA"

Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399

Note: Data from overlapping clone AC090888 [Drafting center: WIBR] was added for finishing. RP11-109D20 spans bases 1-174236 of this sequence. Bases 174237-179373 derive from clone RP11-276K9, AC087790 [Drafting center: WIBR] to establish the overlap with the next BAC in our tiling path. Bases 179373-181312 derive from finished clone CTD-2651B20, AC051619 [Drafting center: UWMS] in order to give overlap required for determining long range contiguity.

FEATURES
 Location/Qualifiers
 1. .181312
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15q15"
 /clones="RP11-109D20"
 /clone_lib="RPCI human BAC library 11"
 /note="This clone overlaps CTD-2014N11 AC090888 and RP11-276K9 AC087790. Data from overlapping BACs were added and the consensus sequence determined from RP11-109D20 to the extent possible."
misc_feature
 1. .89137
 /note="overlap with CTD-2014N11, AC090888"
 36107. .36120
 /note="low quality data"
 40470. .40820
 /note="sequence data generated from subcloned PCR product"
 53030. .53480
 /note="sequence data generated from subcloned PCR product"
 36982. .86983
 /note="low quality data"
 95610. .96030
 /note="sequence data generated from subcloned PCR product"
 96830. .97420
 /note="sequence data generated from subcloned PCR product"
 98260. .98420
 /note="sequence data generated from subcloned PCR product"
 98850. .98852
 /note="single clone coverage"
 100981. .101003
 /note="low quality data"
 174237. .179372
 /note="low quality data"
 100981. .101003
 /note="low quality data"
 174237. .179372
 /note="overlap with RP11-276K9, AC087790. This sequence data was added to establish the overlap between the BACs in the tiling path"
 179373. .181312
 /note="overlap with CTD-2651B20, AC051619. This data is added to give overlap for long range contiguity."

ORIGIN

Query Match 28.1%; Score 405.4; DB 9; Length 181312;
Best Local Similarity 98.6%; Pred. No. 4.8e-75;
Matches 409; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Cy 1 AAGTACGGCTACAGACAGTGAATAATAGTTTCGCTCGCGCTAGAAAACCTCTCTCG 60
 |||||
Db 165016 AAGTACGGCTACAGACAGTGAATAATAGTTTCGCTCGCGCTAGAAAACCTCTCTCG 165075
 |||||

Cy 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGTTCCCTTAACGGAGAGGTGCA 120
 |||||
Db 165076 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGTTCCCTTAACGGAGAGGTGCA 165135
 |||||

Cy 121 GGACTCAGACTTCACAGCCCACTCGGTCCAGCCCTTGACGCAAGAGACGCCAAGGAC 180
 |||||
Db 165136 GGACTCAGACTTCACAGCCCACTCGGTCCAGCCCTTGACGCAAGAGACGCCAAGGAC 165195
 |||||

Cy 181 GCGCTCTCCCGGTCCAGGAGCCCGCAGCTTGCTGGCTGCTGCCGCTCGGTGCAGC 240
 |||||

Db 165196 GCGCTCTCCCGGTCCAGGAGCCCGCAGCTTGCTGGCTGCTGCCGCTCGGTGCAGC 165255
Cy 241 ACTCGCCCGCGTGCAGCATGACCTGTGTGGAACGGCTACTGCTTTTACCCCGAGCCC 300
 |||||
Db 165256 ACTCGCCCGCGTGCAGCATGACCTGTGTGGAACGGCTACTGCTTTTACCCCGAGCCC 165315
 |||||

Cy 301 CCGCATGCCCGCAGGTTTCAGGTTCCACTGCTCATGTTATCTAGTGTGTTTGGCTCTA 360
 |||||
Db 165316 CCGCATGCCCGCAGGTTTCAGGTTCCACTGCTCATGTTATCTAGTGTGTTTGGCTCTA 165375
 |||||

Cy 361 GCAGCAGCTTCCTGCTCATCTTGGCGGGGATCCGTCGCCACTCGCGCTGGTTT 415
 |||||
Db 165376 GCAGCAGCTTCCTGCTCATCTTGGCGGGGATCCGTCGCCACTCGGTAAAGGGTGT 165430
 |||||

RESULT 8
AC009700/c

LOCUS
DEFINITION
 Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
 SEQUENCE, 15 unordered pieces.
ACCESSION
 AC009700
VERSION
 AC009700.4 GI:7622346
KEYWORDS
 HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE
 Homo sapiens (human)
ORGANISM
 Homo sapiens

REFERENCE
AUTHORS
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Lieu, C., Locke, K., McKernan, K., McLaughlin, J., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Mychaleckyj, J., Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
TITLE
 Direct Submission
JOURNAL

REFERENCE
AUTHORS
 1 (bases 1 to 156534)
Homo sapiens chromosome 15, clone RP11-163P10
Unpublished
2 (bases 1 to 156534)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Lieu, C., Locke, K., McKernan, K., McLaughlin, J., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Mychaleckyj, J., Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
TITLE
 Direct Submission
JOURNAL

COMMENT
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 20, 2000 this sequence version replaced gi:6056272.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 163_P_10
Center clone name: L2336
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 108837 bases at least Q40
Consensus quality: 130230 bases at least Q30
Consensus quality: 145227 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 155134; sum-of-contigs
Quality coverage: 3.2 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1169: contig of 1169 bp in length
* 1170 1269: gap of 100 bp
* 1270 3335: contig of 2066 bp in length
* 3336 3435: gap of 100 bp
* 3436 5698: contig of 2263 bp in length
* 5699 5798: gap of 100 bp
* 5799 9216: contig of 3418 bp in length
* 9217 9316: gap of 100 bp
* 9317 15633: contig of 6317 bp in length
* 15634 15733: gap of 100 bp
* 15734 21746: contig of 6013 bp in length
* 21747 21846: gap of 100 bp
* 21847 30351: contig of 8505 bp in length
* 30352 30451: gap of 100 bp
* 30452 36509: contig of 6058 bp in length
* 36510 36609: gap of 100 bp
* 36610 45279: contig of 8670 bp in length
* 45280 45379: gap of 100 bp
* 45380 56851: contig of 11472 bp in length
* 56852 56951: gap of 100 bp
* 56952 67374: contig of 10423 bp in length
* 67375 67474: gap of 100 bp
* 67475 79534: contig of 12060 bp in length
* 79535 79634: gap of 100 bp
* 79635 95432: contig of 15798 bp in length
* 95433 95532: gap of 100 bp
* 95533 120801: contig of 25269 bp in length
* 120802 120901: gap of 100 bp
* 120902 156534: contig of 35633 bp in length.

FEATURES

source

1. 156534
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="15"
/chromosome="15"
/clone="RP11-163P10"
/clone_lib="RPCi-11 Human Male BAC"
1. 1169
/note="assembly_fragment"
1270. 3335
/note="assembly_fragment"
3436. 5698
/note="assembly_fragment"
5799. 9216
/note="assembly_fragment"
9317. 15633
/note="assembly_fragment"
15734. 21746
/note="assembly_fragment"
21847. 30351
/note="assembly_fragment"
30452. 36509
/note="assembly_fragment"
clone_end:17
vector_side:left
36610. 45279
/note="assembly_fragment"
45380. 56851
/note="assembly_fragment"
56952. 67374
/note="assembly_fragment"
67475. 79534
/note="assembly_fragment"

misc_feature 79635. 95432
/note="assembly_fragment"
misc_feature 95533. 120801
/note="assembly_fragment"
misc_feature 120902. 156534
/note="assembly_fragment"

ORIGIN

Query Match 28.0%; Score 403.8; DB 2; Length 156534;
Best Local Similarity 98.3%; Pred. No. 1e-74;
Matches 408; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCCGGTAGAAAACTCTGTGCG 60
Db |||||||
QY 61 GTACCAACCCAGAGCGTTGAG 120
Db |||||||
QY 7088 GTACCAACCCAGAGCGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7029
Db |||||||
QY 121 GGACTCAGACTTCACAG 180
Db |||||||
QY 7028 GGACTCAGACTTCACAG 6969
Db |||||||
QY 181 GCGCTCTCCGCGTCCAG 240
Db |||||||
QY 6968 GCGCTCTCCGCGTCCAG 6909
Db |||||||
QY 241 ACTCGGCGGCGTCCAG 300
Db |||||||
QY 6908 ACTCGGCGGCGTCCAG 6849
Db |||||||
QY 301 CGGCATGCGCGAG 360
Db |||||||
QY 6848 CGGCATGCGCGAG 6789
Db |||||||
QY 361 GCAGCAAGCTTCTGCTCATCTTCCGCGGAGATCCGTCGCGGAGTCCGTCGCGGTTT 415
Db |||||||
QY 6788 GCAGCAAGCTTCTGCTCATCTTCCGCGGAGATCCGTCGCGGAGTCCGTCGCGGTTT 6734
Db |||||||

RESULT 9

AC087790

LOCUS

DEFINITION

AC087790

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC087790 171444 bp DNA linear HTG 09-MAY-2001
Homo sapiens chromosome 15 clone RP11-276K9 map 15, WORKING DRAFT
SEQUENCE, 13 unordered pieces.
AC087790
AC087790.2 GI:13357344
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 171444)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 15, clone RP11-276K9
Unpublished
2 (bases 1 to 171444)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A.,
Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J.,
Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Korton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T.,
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,

RESULT 12
AC009700
LOCUS
DEFINITION Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
AC009700
AC009700.4 GI:7622346
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 156534)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,M., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrim,J., Mollia,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 20, 2000 this sequence version replaced gi:6056272.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 163_P10
Center clone name: 163_P10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 108837 bases at least Q40
Consensus quality: 130230 bases at least Q30
Consensus quality: 145227 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 155134; sum-of-contigs
Quality coverage: 3.2 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1169: contig of 1169 bp in length
* 1170 1269: gap of 100 bp
* 1270 3335: contig of 2066 bp in length
* 3336 3435: gap of 100 bp
* 3436 5699: contig of 2263 bp in length
* 5699 5798: gap of 100 bp
* 5799 9216: contig of 3418 bp in length
* 9217 9316: gap of 100 bp
* 9317 15633: contig of 6317 bp in length
* 15634 15733: gap of 100 bp
* 15734 21746: contig of 6013 bp in length
* 21747 21847: gap of 100 bp
* 21848 30351: contig of 8505 bp in length
* 30352 30451: gap of 100 bp
* 30452 36509: contig of 6058 bp in length
* 36510 45279: gap of 100 bp
* 45280 45379: gap of 100 bp
* 45380 56851: contig of 11472 bp in length
* 56852 56951: gap of 100 bp
* 56952 67374: contig of 10423 bp in length
* 67375 79534: gap of 100 bp
* 79535 79634: contig of 12060 bp in length
* 79635 95432: gap of 100 bp
* 95433 120801: contig of 15798 bp in length
* 120802 120901: gap of 100 bp
* 120902 156534: contig of 25269 bp in length
* 156535 163P10: contig of 35633 bp in length.
FEATURES
Location/Qualifiers
1..156534
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-163P10"
/clone_lib="RPC1-11 Human Male BAC"
1..1169
/note="assembly_fragment"
1270..3335
/note="assembly_fragment"
3436..5698
/note="assembly_fragment"
5799..9216
/note="assembly_fragment"
9317..15633
/note="assembly_fragment"
15734..21746
/note="assembly_fragment"
21847..30351
/note="assembly_fragment"
30452..36509
/note="assembly_fragment"
clone_end:17
vector_side:left
36610..45279
/note="assembly_fragment"
45380..56851
/note="assembly_fragment"
56952..67374
/note="assembly_fragment"
67475..79534
/note="assembly_fragment"
79635..95432
/note="assembly_fragment"
95533..120801
/note="assembly_fragment"
120902..156534
/note="assembly_fragment"
ORIGIN
Query Match 27.5%; Score 396; DB 2; Length 156534;
Best Local Similarity 97.6%; Pred. No. 4.6e-73;
Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1007 TCTGGTACGCTGGCAACGGCGTCCCTGCTGCTCTTCTCGGAGGGCGGTGATC 1066
DB 96111 TTTCGATCCCAACCGCCACAGGCGTCTCTGCTGCTCTTCTCGGAGGGCGGTGATC 96170

QY 1067 TCCAGTATGTTCCGCGCCAGCGCTCTTGGCACCCCTTCTGGACCCAAAGCGCCAGGACTGCA 1126
Db 96171 TCCAGTATGTTCCGCGCCAGCGCTCTTGGCACCCCTTCTGGACCCAAAGCGCCAGGACTGCA 96230
QY 1127 GCCAGGAGAGGGGGCTCACCTCTTATCCTCGGGGACCCACTGACACAGAGCGCGCTC 1186
Db 96231 GCCAGGAGAGGGGGCTCACCTCTTATCCTCGGGGACCCACTGACACAGAGCGCGCTC 96290
QY 1187 TCCAGACTTAAATATGATACCACTAACTCTGTGAGGGGGACCCCAATCTGACTCTCTTC 1246
Db 96291 TCCAGACTTAAATATGATACCACTAACTCTGTGAGGGGGACCCCAATCTGACTCTCTTC 96350
QY 1247 CCGCTTGGGACATCCAGCGCCGGAAGCAGTCCCGCCAGGCTCGGGCCAGGAGCTC 1306
Db 96351 CCGCTTGGGACATCCAGCGCCGGAAGCAGTCCCGCCAGGCTCGGGCCAGGAGCTC 96410
QY 1307 CAGGAAGGGCACTGAGCGCTGTGCGCGGAGGCTCGGACATCCGAGGACCAAGGAAA 1366
Db 96411 CAGGAAGGGCACTGAGCGCTGTGCGCGGAGGCTCGGACATCCGAGGACCAAGGAAA 96470
QY 1367 GTCTCTGGGGCGATCTGTAATAAACCTTTTCTTTTCTTTTAAATAA 1418
Db 96471 GTCTCTGGGGCGATCTGTAATAAACCTTTTCTTTTCTTTTAAATAA 96522
RESULT 13
AX405919
LOCUS AX405919 1636 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 334 from Patent WO20222660.
ACCESSION AX405919
VERSION AX405919.1 GI:21439277
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,
Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 0222660-A 334 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1..1636
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
441..1472
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35026.1"
/db_xref="GI:21439278"
/db_xref="REMBL:CAD35026"
/translation="MATIGHTFPFYAGPKPTFPMDTTLASIIIMFLTALATFIVILPG
IRKTRFLWLLRVVTSFLFIGAAILAVNFSSEWSVQVSTNTSYKAPSEWISADIGLQ
VGLGGVNTLTGTPVQINETINNEEFTWRLGENVAEYAKALEKGLPDPVLYLTK
FTPRSPGLYRQYRLAGHYTSAMLWVAFCLWLLANVMSLPALVYGGYMLLATGIFQL
LALLFPMATSLTSPCLHLGASVLTHGPAFWITLTLLCVLLGLAMAVAHMOP
HRLKAFNQSVDEDPMLNWSPEEGULLSPRYRSMADSPKSDIPLSEASSTKAYCKEA
HPKDPDCAL"

Db 497 GATGACACACCATTGGCCAGCATCATCATGATCTTTCTGACTGCACACTGGCCAGCTTCAT 556
QY 375 GCTCATCTTCCGCGGGATCCGTTGGCCACTCGCGCTGTTTGGTGGTGGAGTCTTCT 434
Db 557 CGTCATCTGCTGCTGCGGAAAGACGAGGCTGTTCTGCTGCTTCCGCTGGTGAC 616
QY 435 CAGTCTGTTATAGCGCGCAAAATTTGTTGCTGTCACCTTCACTGTCAGAAATGTTGTTGG 494
Db 617 CAGCTTATTCATCGGGCTGCAATCCTGGCTGTGAATTTCACTTCTGAGTGTCTGTTGG 676
QY 495 TACAGTGAACACCAACACATCTTACAAAGCCTTCAAGCGAGGCGCTTACAGCCCGTGT 554
Db 677 CCAGTTCAGCACCAACACATCATCAAGGCTTCAAGTCTGAGTGGATCAGGCTGATAT 736
QY 555 CGTCTGCTGCTGCGGCTTGGAGGCAATTAATATACATCACTACAGGACCCAGTGCATCA 614
Db 737 TGGGCTGAGGCTGCGGCTGGGTGGAGTCAACATCACATCACAGGACCCCGCTGCAGCA 796
QY 615 GCTGAACGAGACCAATTAATAACGAGGCTTCACTTGGGCTTCAAGAGGAAATTAACG 674
Db 797 GCTGAATGAGACCAATTAATAACGAGGCTTCACTTGGGCTTGGGTGGGTGAGAACTATGC 856
QY 675 CGCGAGTACCGGACGCACTGAGAGAGGGCTGCGGACCGGCTTCACTTGGGCTGCTCTC 734
Db 857 TGAGGAGTATGCAAGGCTCTGGAGAGGGGCTGCCAGACCTGTGTTGTAACCTAATGA 916
QY 735 GAAGTTCACACCGAGTAGCCCTTGGCGCTTGTACCAACAGTACCACTGGCGGACACTA 794
Db 917 GAAGTTCACCTCAAGAGGCTTGGCTTATACCGGCTTATGCGGAGGCTTCCAGCTGT 976
QY 795 CGCTCGGCGGCTTATGCTGCTGCTTCTGCTTCTGCTTCTTCAACGCTGCTCTC 854
Db 977 CACCTCAGCCATGCTATGCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTC 1036
QY 855 CAGCGCGGCGGCTTCAAGGAGGCTTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 914
Db 1037 CATGCTGCTGCTGCTTATGCTGCTTATGCTGCTTATGCTGCTGCTGCTGCTGCTGCT 1096
QY 915 CG--GGTCTGCTGCTTGGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
Db 1097 GGCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1156
QY 972 CTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
Db 1157 CGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
QY 1032 CTTGCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
Db 1217 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1276
QY 1092 TCGCACCTTCTGACCAAGGCGCAAGGA 1121
Db 1277 GAAGGCTTCTTCAACAGAGTGTGATGA 1306
RESULT 14
BC029819
LOCUS BC029819 1923 bp mRNA linear PRI 06-OCT-2003
DEFINITION Homo sapiens hypothetical protein FLJ32334, mRNA (cDNA clone
MGC:35310 IMAGE:5177619), complete cds.
ACCESSION BC029819
VERSION BC029819.1 GI:20987583
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

BC019755.1 GI:18043594
MGC.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1521)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smillius, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1521)
Strausberg, R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAX Plate: 30 Row: e Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21703789.
Location/Qualifiers
1. 1521
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:25550 IMAGE:3964798"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1521
/gene="BC019755"
/note="synonym: MGC25550"
/db_xref="LocusID:213696"

Db 1085 CATGGCAGTGGCCCA CAGGATGCAGCCTCACAGACTGAAGGCTTCTTCAACCTGAGTTC 1144
Qy 1116 CAAGGACTGCAGCCAGGACAGAGG 1139
Db 1145 AGAGGACCCAGTGTCTGGAGTGGGG 1168

Search completed: June 3, 2004, 11:57:53
Job time : 5819 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 11:58:00 ; Search time 5808 Seconds
(without alignments)
10753.669 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggtacagacag.....aaaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1441	100.0	1441	6	AX035346	AX035346 Sequence
2	976	67.7	1740	9	HSM806241	BX537581 Homo sapi
3	425	29.5	498	6	AX035348	AX035348 Sequence
4	405	28.1	166937	2	AC012255	AC012255 Homo sapi
5	405	28.1	181312	9	AC091117	AC091117 Homo sapi
6	392	27.2	2684	6	AX714511	AX714511 Sequence
7	392	27.2	2684	9	AK056896	AK056896 Homo sapi
8	392	27.2	156534	2	AC009700	AC009700 Homo sapi
9	392	27.2	171444	2	AC087790	AC087790 Homo sapi
10	354	24.6	156534	2	AC009700	AC009700 Homo sapi
11	217	15.1	217	9	HS183H12F	Z57552 H.sapiens C
12	166	11.5	1923	9	BC029819	BC029819 Homo sapi
13	136	9.4	218	9	HS183A12R	Z59954 H.sapiens C
14	39	2.7	1725	9	BC004161	BC004161 Homo sapi
15	39	2.7	1725	9	BC003157	BC003157 Homo sapi
16	38	2.6	1260	9	BC017580	BC017580 Homo sapi
17	38	2.6	2165	6	BD275921	BD275921 Uncouplin
18	38	2.6	2384	6	AX780522	AX780522 Sequence
19	38	2.6	2384	6	AX780523	AX780523 Sequence
20	38	2.6	3993	9	BC036370	BC036370 Homo sapi
21	38	2.6	5642	6	AX345771	AX345771 Sequence
22	38	2.6	10480	6	AX347103	AX347103 Sequence
23	38	2.6	174790	2	AC134964	AC134964 Canis fam
24	38	2.6	180059	2	AC023928	AC023928 Homo sapi
25	38	2.6	180763	2	AC067826	AC067826 Homo sapi
26	38	2.6	185951	2	AC024386	AC024386 Homo sapi
27	38	2.6	186051	2	AC015683	AC015683 Homo sapi
28	37	2.6	1426	10	AF175224	AF175224 Rattus no
29	37	2.6	1640	9	BC050738	BC050738 Homo sapi
30	37	2.6	1678	9	HSM802263	AL137530 Homo sapi
31	37	2.6	1724	9	BC003101	BC003101 Homo sapi
32	37	2.6	2645	3	AY118446	AY118446 Drosophil
33	37	2.6	3734	9	HSM807535	BX647390 Homo sapi
34	37	2.6	3989	9	HSM805415	AL834162 Homo sapi
35	37	2.6	3990	9	HSM805416	AL834352 Homo sapi
36	37	2.6	135628	2	AC067764	AC067764 Homo sapi
37	37	2.6	153166	2	AC048357	AC048357 Homo sapi
38	37	2.6	190371	2	AC116830	AC116830 Mus muscu
39	37	2.6	349980	6	AX344557	AX344557 Sequence
40	36	2.5	1059	3	AF036416	AF036416 Trypanoso
41	36	2.5	1407	10	BC013450	BC013450 Mus muscu
42	36	2.5	1603	10	BC024516	BC024516 Mus muscu
43	36	2.5	1713	9	BC003624	BC003624 Homo sapi
44	36	2.5	2155	9	BC015749	BC015749 Homo sapi
45	36	2.5	2255	6	AR168373	AR168373 Sequence

ALIGNMENTS

RESULT 1	AX035346	AX035346	Sequence 1	1441 bp	DNA	linear	PAT 15-NOV-2000
LOCUS	AX035346	Sequence 1	from Patent WO053748.				
DEFINITION	AX035346	AX035346					
ACCESSION	AX035346	AX035346.1	GI:11191064				
VERSION	AX035346.1	GI:11191064					
KEYWORDS		Homo sapiens (human)					
SOURCE		Homo sapiens					
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1	Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.					
AUTHORS		Novel compounds					
TITLE		Patent: WO 0053748-A 1 14-SEP-2000;					
JOURNAL							

```
BRUCK CLAUDINE ELVIRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ;
VINALS Y DE BASSOLS CARLOTA (BE) ; COCHE THIERRY (BE) ; CASSART
JEAN POL (BE)
Location/Qualifiers
1. .1441
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

```

h      100.0%; Score 1441; DB 6; Length 1441;
Similarity 100.0%; Pred.No. 0;
4i; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1	AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCCGCTAGAAAACTCTGTGC	60
1	AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCCGCTAGAAAACTCTGTGC	60
1	GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCCTTAAACGAGAGGTGCA	120
1	GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCCTTAAACGAGAGGTGCA	120
1	GGACTCAGACTTTCACAGCCCCACTCGGTCCCAGCCTTGTAACGAAAAGAGACGCAAGGAC	180
1	GGACTCAGACTTTCACAGCCCCACTCGGTCCCAGCCTTGTAACGAAAAGAGACGCAAGGAC	180
1	GGCTCTCCCGGTCACAGCCCACTCGGTCCCAGCCTTGCTGGCTTGCTGCCCGCTGGTGCA	240
1	GGCTCTCCCGGTCACAGCCCACTCGGTCCCAGCCTTGCTGGCTTGCTGCCCGCTGGTGCA	240
1	ACTCGGCGGGGTGCAGCATGAACCTGTGGAAACGGGCTACTGCCTTTTACCCAGCC	300
1	ACTCGGCGGGGTGCAGCATGAACCTGTGGAAACGGGCTACTGCCTTTTACCCAGCC	300
1	CGGCATGCGCAGGCTTCAGCGTTCCACTGCTCATCGTTATTCAGTGTTCCTTGGCTCTA	360
1	CGGCATGCGCAGGCTTCAGCGTTCCACTGCTCATCGTTATTCAGTGTTCCTTGGCTCTA	360
1	GCAGCAAGCTTCCTGCTCATCTTGC CGGGATCCGTTGGCCACTCGCGCTGTTTGGTTG	420
1	GCAGCAAGCTTCCTGCTCATCTTGC CGGGATCCGTTGGCCACTCGCGCTGTTTGGTTG	420
1	GTGAGAGTTCCTTCAGTCTGTTTCATAGGCGCAGAAATGTGGCTGTGCACTTCAGTGCA	480
1	GTGAGAGTTCCTTCAGTCTGTTTCATAGGCGCAGAAATGTGGCTGTGCACTTCAGTGCA	480
1	GAATGGTTTCGTTGGGTACAGTGAAACCAACACATCTCACAAAGCCTTCAGCGCAGCGCG	540
1	GAATGGTTTCGTTGGGTACAGTGAAACCAACACATCTCACAAAGCCTTCAGCGCAGCGCG	540
1	GTACAGCCCGTGTGCGTCTGCTCGTGGGCTGGAGGCAATTAATTAACACTCACAGGG	600
1	GTACAGCCCGTGTGCGTCTGCTCGTGGGCTGGAGGCAATTAATTAACACTCACAGGG	600
1	ACCCAGTGCATCAGCTGAACGAGACCATTGACTACAAACGAGCAGTTTACCTGGCGTCTG	660
1	ACCCAGTGCATCAGCTGAACGAGACCATTGACTACAAACGAGCAGTTTACCTGGCGTCTG	660
1	AAAGAGAAATTAACGCGGAGTACGGAACGCACTGGAAGAAGGGCTGCCGACCCAGTG	720
1	AAAGAGAAATTAACGCGGAGTACGGAACGCACTGGAAGAAGGGCTGCCGACCCAGTG	720
1	CTCTACCTGCGGAGAAATTCACACCGAGTAGCCCTTGGCGCCTGTACCAACCACTACCAC	780
1	CTCTACCTGCGGAGAAATTCACACCGAGTAGCCCTTGGCGCCTGTACCAACCACTACCAC	780
1	CTGGCGGACACTACGCCCTCGGCCACCGCTATGGGTGGCTTCTGCTTCTGGCTCTCTCC	840
1	CTGGCGGACACTACGCCCTCGGCCACCGCTATGGGTGGCTTCTGCTTCTGGCTCTCTCC	840
1	AACGTGCTGCTCTCCACGCGGGCCCCGCTCTACGGAGCTTGGCACTGTGACCAACCGGA	900
1	AACGTGCTGCTCTCCACGCGGGCCCCGCTCTACGGAGCTTGGCACTGTGACCAACCGGA	900

901	QY	GCCTTCGGCGCTCTTCGGGGTCTTCGGCTTGGCCCTCCATCTCTAGCGTGCCTCTGCGCG	960
901	Db	GCCTTCGGCGCTCTTCGGGGTCTTCGGCTTGGCCCTCCATCTCTAGCGTGCCTCTGCGCG	960
961	QY	CTCGGCCTAGGCTCTCCCGGCTCACCACCTACGTACGGCGCGGCTTCTGGGTCAAGCTG	1020
961	Db	CTCGGCCTAGGCTCTCCCGGCTCACCACCTACGTACGGCGCGGCTTCTGGGTCAAGCTG	1020
1021	QY	GCAACCGGCGTCTCTGTGCTCTTCTCGGAGGGCGCTGGTGAGTCTCCAGTATGTTCTGG	1080
1021	Db	GCAACCGGCGTCTCTGTGCTCTTCTCGGAGGGCGCTGGTGAGTCTCCAGTATGTTCTGG	1080
1081	QY	CCAGCGCTCTTCGCACCCCTTCCTGGACCCAAAGGCCAAGGACTGCAGCCAGGAGAGGG	1140
1081	Db	CCAGCGCTCTTCGCACCCCTTCCTGGACCCAAAGGCCAAGGACTGCAGCCAGGAGAGGG	1140
1141	QY	GGCTCACCTCTTATCTCTCGGCGACCCCACTGCACAGCAGGCGCTCTCCAGACTTAAAA	1200
1141	Db	GGCTCACCTCTTATCTCTCGGCGACCCCACTGCACAGCAGGCGCTCTCCAGACTTAAAA	1200
1201	QY	TGTTATCACCACTAACTGTGAGGGGACCCCAATCTGGACTCTCTTCCCGCCTTGGGACAT	1260
1201	Db	TGTTATCACCACTAACTGTGAGGGGACCCCAATCTGGACTCTCTTCCCGCCTTGGGACAT	1260
1261	QY	CGCAGGCCGGGAAGCAGTGCCTCGCAGGCGCTTGGGCCAGGAGAGCTTCAGGAAGGGCACTG	1320
1261	Db	CGCAGGCCGGGAAGCAGTGCCTCGCAGGCGCTTGGGCCAGGAGAGCTTCAGGAAGGGCACTG	1320
1321	QY	AGCGCTGTGGCGGAGGCCTCGGACATCCGACGGCACCGGGAAGTCTCCTGGGGCGA	1380
1321	Db	AGCGCTGTGGCGGAGGCCTCGGACATCCGACGGCACCGGGAAGTCTCCTGGGGCGA	1380
1381	QY	TCTGTAAATAAACCTTTTTTTTCTTTTGTTTTTTAAAAAATAAAAAAAAAAAAAA	1440
1381	Db	TCTGTAAATAAACCTTTTTTTTCTTTTGTTTTTTAAAAAATAAAAAAAAAAAAAA	1440
1441	QY	A 1441	
1441	Db	A 1441	

```

RESULT 2
HSM806241
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFZp686C04213 (from clone DKFZp686C04213).
ACCESSION BX537581
VERSION BX537581.1 GI:31873659
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1740)
Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
Direct Submission
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Netherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686C04213) is available at the RZPD in Berlin.
please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1..1740
/organism="Homo sapiens"
/mol_type="mRNA"
FEATURES
source

```



```

/misc_feature /note="sequence data generated from subcloned PCR product"
53030. .53480
/note="sequence data generated from subcloned PCR product"
86982. .86983
/note="low quality data"
95610. .96030
/note="sequence data generated from subcloned PCR product"
96830. .97420
/note="sequence data generated from subcloned PCR product"
98260. .98420
/note="single clone coverage"
98850. .98852
/note="low quality data"
100981. .101003
/note="low quality data"
174237. .179372
/note="overlap with RP11-276X9, AC087790. This sequence
data was added to establish the overlap between the BACs
in the tiling path"
179373. .181312
/note="overlap with CTE-2651B20, AC051619. This data is
added to give overlap for long range contiguity."

```

ORIGIN

Query Match	28.1%;	Score 405;	DB 9;	Length 181312;
Best Local Similarity	100.0%;	Pred. No. 7.4e-213;		
Matches 405;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AAAGTAA	CGGCTACAGACAGT	GAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTGC 60
Db	165016	AAAGTAA	CGGCTACAGACAGT	GAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTGC 165075
QY	61	GTACCAA	CCCCAGAGCGTTG	AGAGCAGCCCACTCCACGCTTCCTTAA
Db	165076	GTACCAA	CCCCAGAGCGTTG	AGAGCAGCCCACTCCACGCTTCCTTAA
QY	121	GGACTCAGACTT	CACCGCCCACTCGGT	CCCCAGCCTTGTACGCAAGAGACGCGCAAGGAC 180
Db	165136	GGACTCAGACTT	CACCGCCCACTCGGT	CCCCAGCCTTGTACGCAAGAGACGCGCAAGGAC 165195
QY	181	GCGCTCT	CCCCGGTCCAGG	CAGCCCCAGCTTGTGTGGCTTGGCTGCCCGCTGGTGCAGC 240
Db	165196	GCGCTCT	CCCCGGTCCAGG	CAGCCCCAGCTTGTGTGGCTTGGCTGCCCGCTGGTGCAGC 165255
QY	241	ACTCGGCCGGGT	CAGCATGACCTGT	GGAACGGCGTACTGCTTTTACCCCCAGCCC 300
Db	165256	ACTCGGCCGGGT	CAGCATGACCTGT	GGAACGGCGTACTGCTTTTACCCCCAGCCC 165315
QY	301	CGGCATG	CCGCGAGGCTT	CAGCGTTCACATGCTCATCGTTATTCATGTGTTTGGCTCTA 360
Db	165316	CGGCATG	CCGCGAGGCTT	CAGCGTTCACATGCTCATCGTTATTCATGTGTTTGGCTCTA 165375
QY	361	GCAGCAAGCTT	CTCTGCTCATCTT	CGCGGGGATCCGTGGCCACTCG 405
Db	165376	GCAGCAAGCTT	CTCTGCTCATCTT	CGCGGGGATCCGTGGCCACTCG 165420

RESULT 6

AX714511/c	AX714511	2684 bp	DNA	linear	PAT 15-APR-2003
LOCUS	Sequence	1195 from Patent EP1293569.			
DEFINITION	AX714511				
ACCESSION	AX714511.1	GI:29889464			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.				

```

TITLE
JOURNAL
Full-length cDNAs
Patent: EP 1293569-A 1195 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)
FEATURES
Location/Qualifiers
source
1..2684
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

```

Query Match	27.2%;	Score 392;	DB 6;	Length 2684;
Best Local Similarity	100.0%;	Pred. No. 1.2e-205;		
Matches 392;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1027	GGCGTCGTGTGCCTCTTCTCGGAGGGCCGTGTGAGTCTCCAGTATGTTCCGCCACG	1086		
DB				
2343	GGCGTCGTGTGCCTCTTCTCGGAGGGCCGTGTGAGTCTCCAGTATGTTCCGCCACG	2284		
QY 1087	GCTCTTCGACCCCTTCTTCGACCCAAAGCGCCAAAGCACTGCAGCCAGGAGAGAGGGGGCTCA	1146		
DB				
2283	GCTCTTCGACCCCTTCTTCGACCCAAAGCGCCAAAGCACTGCAGCCAGGAGAGAGGGGGCTCA	2224		
QY 1147	CCTCTTATCTCGGCGACCCACTGCAACAGCAGGCCGCTCTCTCCAGACTTAAATGTATC	1206		
DB				
2223	CCTCTTATCTCGGCGACCCACTGCAACAGCAGGCCGCTCTCTCCAGACTTAAATGTATC	2164		
QY 1207	ACCCTAACCTGTGAGGGGACCCCAATCTGGACTCCTTCCCCGCCCTTGGGACATCGCAGG	1266		
DB				
2163	ACCCTAACCTGTGAGGGGACCCCAATCTGGACTCCTTCCCCGCCCTTGGGACATCGCAGG	2104		
QY 1267	CCGGGAAGCAGTGCCCGCCAGGCCCTGGGCCAGGAGAGCTCCAGGAAGGGCACTGAGCGCT	1326		
DB				
2103	CCGGGAAGCAGTGCCCGCCAGGCCCTGGGCCAGGAGAGCTCCAGGAAGGGCACTGAGCGCT	2044		
QY 1327	GCTGGCGGAGGCCCTCGGACATCCGCGAGCACCCAGGGAAAGTCTCCTGGGCGGATCTGTA	1386		
DB				
2043	GCTGGCGGAGGCCCTCGGACATCCGCGAGCACCCAGGGAAAGTCTCCTGGGCGGATCTGTA	1984		
QY 1387	AATAAACCTTTTTTCTTTTGTTTTTTAAAAA	1418		
DB				
1983	AATAAACCTTTTTTCTTTTGTTTTTTAAAAA	1952		

RESULT 7

AK056896	2684 bp	mRNA	linear	PRI 01-AUG-2002
LOCUS				
DEFINITION	Homo sapiens cDNA FLJ32334 fis, clone PROST2005426.			
ACCESSION	AK056896			
VERSION	AK056896.1 GI:16552419			
KEYWORDS	oligo capping; fis (full insert sequence).			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1			
AUTHORS	Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,			

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1. The Effect of Temperature on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	Smith, J. D.	2. The Effect of Concentration on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education
3. The Effect of pH on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education	1965	Smith, J. D.	4. The Effect of Catalyst on the Rate of Reaction of Hydrogen Peroxide with Potassium Iodide	Journal of Chemical Education

Nishiyama, K., Masuno, I., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2684)
 Isogai, T., Otsuki, T. and Sugiyama, T.
 Direct Submission
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source Location/Qualifiers
1 .2684
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PROST2005426"
/tissue_type="prostate"
/clone_lib="PROST2"
/note="cloning vector: pME18SFL3"
56 .952
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71304.1"
/db_xref="GI:16552420"
/translation="MATLGHFPFYAGKPTFPMDTTLASIMFLTALATFIVILPG
IRKLPFLVLRVYVTSIFGAAILGTPVQQLNETINIEFTWRLGANYAEYAKALE
KGLPDPVLYLAEKFTPRSPGLYRQYRLAGHYTSAMLVAFCLWLLGANYLMPVLVY
GGYMLLATGIFOLLALLFESMATSLTSPCLHLGASVLHHTHGPFAFWLTTGLLCLVL
LGLAMAVAHMQPHRLKARFNQSVDEDPLEWSPREGGLSPRYRSMADSPKSDIPL
SEASSTKAYCKEAPHKOPDCAL"

ORIGIN

Query Match 27.2% Score 392; DB 9; Length 2684;
Best Local Similarity 100.0%; Pred. No. 1.2e-205;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1027 GCGTCTCTGTCCTCTCTCTCGGAGGCGCGTGGTGAGTCTCCAGTATGTTCCGGCCAGC 1086
Db 2343 GCGTCTCTGTCCTCTCTCTCGGAGGCGCGTGGTGAGTCTCCAGTATGTTCCGGCCAGC 2284
Qy 1087 GCTCTTCGACCCCTTCTGGACCAAGCGCCCAAGGACTGCGAGAGAGAGGGGCTCA 1146
Db 2283 GCTCTTCGACCCCTTCTGGACCAAGCGCCCAAGGACTGCGAGAGAGAGGGGCTCA 2224
Qy 1147 CCTCTATCTCTCGGACCCACTGCACAAAGCAGCGCGCTCTCCAGACTTAAATGTATC 1206
Db 2223 CCTCTATCTCTCGGACCCACTGCACAAAGCAGCGCGCTCTCCAGACTTAAATGTATC 2164
Qy 1207 ACCACTAACCTGTAGGGGACCCCAATCTGGACTCTCTCCCGCTTGGGACATCGCAGG 1266
Db 2163 ACCACTAACCTGTAGGGGACCCCAATCTGGACTCTCTCCCGCTTGGGACATCGCAGG 2104
Qy 1267 CCGGAACAGTGCCTCGGAGGCTTGGCCAGGAGAGCTCCAGGAAGGCACTGAGCGCT 1326
Db 2103 CCGGAACAGTGCCTCGGAGGCTTGGCCAGGAGAGCTCCAGGAAGGCACTGAGCGCT 2044
Qy 1327 GCTGGCGGAGGCTCGGACATCGGAGCACCAGGGAAGTCTCTCTGGGCGGATCTGTA 1386
Db 2043 GCTGGCGGAGGCTCGGACATCGGAGCACCAGGGAAGTCTCTCTGGGCGGATCTGTA 1984
Qy 1387 AATAAACCTTTTCTTTTCTTTTAAAAA 1418
Db 1983 AATAAACCTTTTCTTTTCTTTTAAAAA 1952

RESULT 8
AC009700 156534 bp DNA linear HTG 20-APR-2000
LOCUS
DEFINITION Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
AC009700
ACCESSION AC009700.4 GI:7622346
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 156534)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP11-163P10
Unpublished
2 (bases 1 to 156534)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Depayre,B., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Hearford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 20, 2000 this sequence version replaced gi:6056272.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2336
Center clone name: 163_P10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 108837 bases at least Q40
Consensus quality: 130230 bases at least Q30
Consensus quality: 145227 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 155134; sum-of-contigs
Quality coverage: 3.2 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1169: contig of 1169 bp in length
* 1170 1269: gap of 100 bp
* 1270 3335: contig of 2066 bp in length
* 3336 3435: gap of 100 bp
* 3436 5698: contig of 2263 bp in length
* 5699 5798: gap of 100 bp
* 5799 9216: contig of 3418 bp in length
* 9217 9317: gap of 100 bp
* 9317 15633: contig of 6317 bp in length
* 15634 15733: gap of 100 bp
* 15734 21746: contig of 6013 bp in length
* 21747 21846: gap of 100 bp
* 21847 30351: contig of 8505 bp in length
* 30352 30451: gap of 100 bp
* 30452 36509: contig of 6058 bp in length
* 36510 36609: gap of 100 bp

DLPKWDYRRDARAFPHLEPTPCVTSRHTPLI"

ORIGIN

Query Match 11.5%; Score 166; DB 9; Length 1923;
Best Local Similarity 100.0%; Pred. No. 2.7e-80;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1216 CTGTGAGGGGACCCAAATCTGGACTCCTTCCCGCCTTGGACATCGCAGCGCGGAGC 1275
|||||
Db 1557 CTGTGAGGGGACCCAAATCTGGACTCCTTCCCGCCTTGGACATCGCAGCGCGGAGC 1498
|||||
QY 1276 AGTCCCGCCAGGCTGGCCAGGAGAGCTCCAGGAAGGCACTGAGCGCTGCTGGCGCG 1335
|||||
Db 1497 AGTCCCGCCAGGCTGGCCAGGAGAGCTCCAGGAAGGCACTGAGCGCTGCTGGCGCG 1438
|||||
QY 1336 AGCCCTGGACATCCGAGGACCAGGAAAGTCTCTCTGGGCGGAT 1381
|||||
Db 1437 AGCCCTGGACATCCGAGGACCAGGAAAGTCTCTCTGGGCGGAT 1392
|||||

RESULT 13

HS183A12R 218 bp DNA linear PRI 19-OCT-1995
LOCUS H.sapiens CpG island DNA genomic MseI fragment, clone 183a12,
DEFINITION reverse read cpgl83a12.r1a.
ACCESSION Z59954
VERSION Z59954.1 GI:1031867
KEYWORDS CpG island; genomic MseI fragment.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Cross, S.H., Charlton, J.A., Nan, X. and Bird, A.P.
AUTHORS Purification of CpG islands using a methylated DNA binding column
TITLE Nat. Genet. 6 (3), 236-244 (1994)
JOURNAL 94282070
MEDLINE
PUBMED 8012384
REFERENCE 2 (bases 1 to 218)
AUTHORS Dodsworth, S.J., Huckle, E., Wilkinson, P. and Micklem, G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT Vector: pGEM-5Zf(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES

source
1. .218
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="183a12"
/sex="male"
/tissue_type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"

ORIGIN

Query Match 9.4%; Score 136; DB 9; Length 218;
Best Local Similarity 100.0%; Pred. No. 1.2e-63;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1268 CGGGAAGCAGTCCCGCAGGCTGGCCAGGAGAGTCCAGGAAGGCACTGAGCGCTG 1327
|||||
Db 72 CGGGAAGCAGTCCCGCAGGCTGGCCAGGAGAGTCCAGGAAGGCACTGAGCGCTG 131
|||||
QY 1328 CTGGCGCGAGGCTCGGACATCCGAGGACCAGGGAAGTCTCTCTGGGCGGATCTGTAA 1387
|||||
Db 132 CTGGCGCGAGGCTCGGACATCCGAGGACCAGGGAAGTCTCTCTGGGCGGATCTGTAA 191
|||||
QY 1388 ATAAACCTTTTTTCT 1403

|||||
192 ATAAACCTTTTTTCT 207

Db

RESULT 14

BC004161 1725 bp mRNA linear PRI 12-NOV-2003
LOCUS Homo sapiens transmembrane 4 superfamily member tetraspan NET-7,
DEFINITION mRNA (cdna clone MGC:2447 IMAGE:2958221), complete cds.
ACCESSION BC004161
VERSION BC004161.2 GI:33872773
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1725)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1725)
Strausberg, R.
Direct Submission
Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:13278776.
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amad@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

REMARK

COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 2 Row: k Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21264576.

FEATURES

source
1. .1725
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:2447 IMAGE:2958221"
/tissue_type="kidney, renal cell adenocarcinoma"
/clone_lib="NIH_MGC_14"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 13:05:55 ; Search time 4062 Seconds
(without alignments)
10593.649 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	528	36.6	537	9 AI821606	AI821606 nk08a11.x
C 2	515	35.7	515	9 AI826629	AI826629 wk35e04.x
C 3	510	35.4	510	9 AI304327	AI304327 qo57h06.x
C 4	446	31.0	446	9 AI984141	AI984141 wu21c02.x

C 5	30.1	434	9	AI991272	AI991272 wu41h04.x
C 6	28.7	414	9	AA593860	AA593860 nn19f03.s
C 7	28.1	405	9	AI274929	AI274929 q149c11.x
C 8	27.0	505	9	AI791844	AI791844 nk08a11.y
C 9	26.9	388	9	AI660493	AI660493 we67h02.x
C 10	26.5	411	9	AI983793	AI983793 wu20c09.x
C 11	26.4	381	9	AI281211	AI281211 qk58e08.x
C 12	26.0	573	14	CB854121	CB854121 UI-CF-DU1
C 13	26.0	405	9	AI346155	AI346155 qp43f12.x
C 14	25.6	486	9	AI660560	AI660560 we68b08.x
C 15	24.1	737	14	CB305399	CB305399 UI-CF-EN1
C 16	22.9	466	9	AA573825	AA573825 nk08a11.s
C 17	22.0	324	12	BM987789	BM987789 UI-H-CO0-
C 18	20.5	296	9	AI821178	AI821178 ne17a03.y
C 19	20.2	348	9	AW050605	AW050605 wz19b11.x
C 20	19.4	690	12	BM977010	BM977010 UI-CF-EN1
C 21	19.2	277	9	AI732165	AI732165 ne17a03.x
C 22	18.7	687	14	CB850544	CB850544 UI-CF-EN1
C 23	18.0	260	9	AW009962	AW009962 ws89h06.x
C 24	17.5	252	9	AI262416	AI262416 qk38e04.x
C 25	12.8	539	9	AI924216	AI924216 wn03b10.x
C 26	12.7	593	12	BM973444	BM973444 UI-CF-EC1
C 27	11.5	315	9	AA618335	AA618335 nq15g11.s
C 28	10.3	274	9	AA469031	AA469031 ne17a03.s
C 29	6.5	602	12	BQ017315	BQ017315 UI-H-DT1-
C 30	6.1	151	9	AI695625	AI695625 we50h04.x
C 31	6.1	378	10	BF917041	BF917041 IL3-UT011
C 32	4.4	119	9	AI921289	AI921289 wo22h06.x
C 33	4.2	942	14	CF241985	CF241985 AGENCOURT
C 34	4.1	279	9	AI270429	AI270429 qu84h03.x
C 35	2.8	175	14	CF316906	CF316906 HD--06-G1
C 36	2.8	217	14	CF328864	CF328864 NACL--03-
C 37	2.7	188	12	BM887036	BM887036 sam33c08.
C 38	2.7	225	9	AI932794	AI932794 wo36g03.x
C 39	2.7	243	13	BQ394290	BQ394290 NISC_rg08
C 40	2.7	248	9	AI537827	AI537827 tp35h10.x
C 41	2.7	250	14	CD421839	CD421839 laa66d09.
C 42	2.7	284	9	AI537819	AI537819 tp35g10.x
C 43	2.7	300	9	AI475331	AI475331 tl81h09.x
C 44	2.7	324	10	AW130187	AW130187 xf29f09.x
C 45	2.7	328	13	BU055639	BU055639 UI-M-F00-

ALIGNMENTS

RESULT 1
AI821606/c
LOCUS
DEFINITION
AI821606 537 bp mRNA linear EST 13-DEC-1999
nk08a11.x5 NCI CGAP Co2 Homo sapiens cDNA clone IMAGE:1012892 3',
similar to contains_TARI.t3 TARI MER22 repetitive element ;, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AI821606
nk08a11.x5
similar to contains_TARI.t3
sequence.
AI821606
AI821606.1
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 537)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other_ESTs: nk08a11.y5
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

Qy	1262	GCAGGCCGGGAAGCAGTGCCTCCGCCCAGGCCCTGGGCCAGGAGAGCTCCAGGAAGGGCACTGA	1321
Db	155	GCAGGCCGGGAAGCAGTGCCTCCGCCCAGGCCCTGGGCCAGGAGAGCTCCAGGAAGGGCACTGA	96
Qy	1322	GCGCTGCTGGCGGAGGCGCTCGGACATCCGAGGGCACCCAGGAAAGTCTCCTGGGGCGAT	1381
Db	95	GCGCTGCTGGCGGAGGCGCTCGGACATCCGAGGGCACCCAGGAAAGTCTCCTGGGGCGAT	36
Qy	1382	CTGTAAATAAAACCTTTTTTCTTTTGTTTTTTAAA	1416
Db	35	CTGTAAATAAAACCTTTTTTCTTTTGTTTTTTAAA	1

RESULT 3					
AI304327/c					
LOCUS	AI304327	510 bp	mRNA	linear	EST 01-FEB-1999
DEFINITION	Q057h06.xl NCI_CGAP C08 Homo sapiens cDNA clone IMAGE:1912667 3'				
	similar to contains TAR1.t3 TARI repetitive element ; mRNA sequence.				

ACCESSION AI304327
 VERSION AI304327.1 GI:3988016
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 510)
 REFERENCE NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1487 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 476.

```

FEATURES
source
Location/Qualifiers
1. .510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1912667"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP C08"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

```

```

ORIGIN
Query Match 35.4%; Score 510; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. NO. 5.2e-92;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy	904	TTCGCGCTCTTGGGGTCTTCGCCTTGCCCTCCATCTCTAGCGTCCGCTCTGCCCGCTC	963
Dbb	510	TTCGCGCTCTTGGGGTCTTCGCCTTGCCCTCCATCTCTAGCGTCCGCTCTGCCCGCTC	451
Qy	964	CGCCTAGGTCTCCTCGCGCTCACCACTAGTA CGGCGCCGCTTCTTGGGTACGCTGGCA	1023

Db	450	CGCCTAGGCTCCTCCGGCTCACCACTCAGTACGGCGCGCCTTCTGGGTACGCTGGCA	391
Qy	1024	ACCGGCGTCTGTGCTTCTTCTCCGAGGGGCGTGTGAGTCTCCAGTATGTTCCGSCC	1083
Db	390	ACCGGCGTCTGTGCTTCTTCTCCGAGGGGCGTGTGAGTCTCCAGTATGTTCCGSCC	331
Qy	1084	AGCGCTCTTCGCA CCGTCTCTGGACCAAGCGCCCAAGGACTGCAGCCAGAGAGAGGGGC	1143
Db	330	AGCGCTCTTCGCA CCGTCTCTGGACCAAGCGCCCAAGGACTGCAGCCAGAGAGAGGGGC	271
Qy	1144	TCACCTCTTATCCTCGGCGACCCACTGCACAAGCGGCGCTCTCCAGACTTTAAATGT	1203
Db	270	TCACCTCTTATCCTCGGCGACCCACTGCACAAGCGGCGCTCTCCAGACTTTAAATGT	211
Qy	1204	ATCACCACTAACCTGTAGGGGGA CCGAATCTGGACTCTCTCCCGCTTGGGACATCGC	1263
Db	210	ATCACCACTAACCTGTAGGGGGA CCGAATCTGGACTCTCTCCCGCTTGGGACATCGC	151
Qy	1264	AGGCGGGAAGCAGTGC CCGCCAGGCTGGCCAGAGAGCTCCAGGAAGGCACCTGAGC	1323
Db	150	AGGCGGGAAGCAGTGC CCGCCAGGCTGGCCAGAGAGCTCCAGGAAGGCACCTGAGC	91
Qy	1324	GCTGCTGSCGCGAGGCTTCGGACATCCG CAGGCAC CAGGGGAAAGTCTCTGGGGCGATCT	1383
Db	90	GCTGCTGSCGCGAGGCTTCGGACATCCG CAGGCAC CAGGGGAAAGTCTCTGGGGCGATCT	31
Qy	1384	GTAAATAAAACCTTTTTTCTTTTGTTTTTT	1413
Db	30	GTAAATAAAACCTTTTTTCTTTTGTTTTTT	1

RESULT 4	AI984141/c	AI984141	446 bp	mRNA	linear	EST 27-OCT-1999
LOCUS		wu21c02.x1 Soares Dieckgraefe Colon_NHCD Homo sapiens cDNA clone				
DEFINITION		IMAGE:2520674 3', mRNA sequence.				
ACCESSION		AI984141				
VERSION		AI984141.1	GI:5811360			
KEYWORDS		EST.				
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE		1 (bases 1 to 446)				
AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .				
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
		Tumor Gene Index				
JOURNAL		Unpublished (1997)				
COMMENT		Contact: Robert Strausberg, Ph.D.				

```

seq primer: 5'400P from GIBCO.
Location/Qualifiers
1..446
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2520674"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares Dieckgraefe colon NHCD"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,

```


JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1204 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 406.
Location/Qualifiers

FEATURES

source
1. 414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1084349"
/sex="mixed"
/tissue_type="colon tumor"
/lab_host="NCI CGAP Col2"
/note="Organ: colon; Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled colon tumors. 5' adaptor sequence: 5' GAATTCGGACGAG 3' 3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' Average insert size: 1.2 kb."

ORIGIN

Query Match 28.7%; Score 414; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 6.7e-73;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 999 CGCGGCTTCTGGTACGCTGGCAACCGGCGTCTGTGCTCTTCTCGGAGGGCGGT 1058
Db 414 CGCGGCTTCTGGTACGCTGGCAACCGGCGTCTGTGCTCTTCTCGGAGGGCGGT 355
QY 1059 GGTGAGTCTCCAGTATGTTCCGCCCGGCGTCTTGGCACCTTCTGGACCAAGCGCCAA 1118
Db 354 GGTGAGTCTCCAGTATGTTCCGCCCGGCGTCTTGGCACCTTCTGGACCAAGCGCCAA 295
QY 1119 GGACTGCAGCCAGGAGAGAGGGGCTCACCTCTTATCTCGGAGCCCACTGCACAAGCA 1178
Db 294 GGACTGCAGCCAGGAGAGAGGGGCTCACCTCTTATCTCGGAGCCCACTGCACAAGCA 235
QY 1179 GGCGCTCTCCAGACTTAAATGATATCACTTAACTGAGGGGAGACCAATCTGGA 1238
Db 234 GGCGCTCTCCAGACTTAAATGATATCACTTAACTGAGGGGAGACCAATCTGGA 175
QY 1239 CTCCTTCCCGCTTGGGACATCGCAGGCGGGAGAGCAGTGCCCGCCAGGCTGGGCCAG 1298
Db 174 CTCCTTCCCGCTTGGGACATCGCAGGCGGGAGAGCAGTGCCCGCCAGGCTGGGCCAG 115
QY 1299 GAGAGCTCCAGGAGGGCACTGAGCGCTGTGGCGCGAGGCTCGGACATCCGAGGCAC 1358
Db 114 GAGAGCTCCAGGAGGGCACTGAGCGCTGTGGCGCGAGGCTCGGACATCCGAGGCAC 55
QY 1359 CAGGAAAGTCTCTGGGGCGATCTGTAAATAAACCTTTTCTTTTCTTTT 1412
Db 54 CAGGAAAGTCTCTGGGGCGATCTGTAAATAAACCTTTTCTTTTCTTTT 1

RESULT 7

AI274929/c 405 bp mRNA linear EST 29-JAN-1999
LOCUS
DEFINITION
ql49c11.x1 NCI CGAP_C08 Homo sapiens CDNA clone IMAGE:1875668 3'
similar to contains TARI.t3 TARI repetitive element 1, mRNA
sequence.
ACCESSION
AI274929
VERSION
AI274929.1 GI:3897203
KEYWORDS
EST.

SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 405)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1458 Std Error: 0.00
Seq primer: -40Up from Gibco
High quality sequence stop: 395.
Location/Qualifiers

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1458 Std Error: 0.00
Seq primer: -40Up from Gibco
High quality sequence stop: 395.
Location/Qualifiers

FEATURES

source
1. 405
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1875668"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP_C08"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 28.1%; Score 405; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 4.1e-71;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1009 TGGTCACTGCGCAACCGGCGTCTGTGCTCTTCTCGAGGGCGGTGAGTCTC 1068
Db 405 TGGTCACTGCGCAACCGGCGTCTGTGCTCTTCTCGAGGGCGGTGAGTCTC 346
QY 1069 CAGTATGTTCCGCCCGGCGTCTTGGCACCTTCTGGACCAAGCGCCAGGACTGCAGC 1128
Db 345 CAGTATGTTCCGCCCGGCGTCTTGGCACCTTCTGGACCAAGCGCCAGGACTGCAGC 286
QY 1129 CAGGAGAGAGGGGCTCACCTCTTATCTCGGCGAGCCCACTGCACAAGCGCCGCTCTC 1188
Db 285 CAGGAGAGAGGGGCTCACCTCTTATCTCGGCGAGCCCACTGCACAAGCGCCGCTCTC 226
QY 1189 CCAGACTTAAATGATATCACTAACTTGTGAGGGGAGCCCAATCTGGACTCTTCCCC 1248
Db 225 CCAGACTTAAATGATATCACTAACTTGTGAGGGGAGCCCAATCTGGACTCTTCCCC 166
QY 1249 GCCTTGGGACATCGCAGGCGGGAGAGCAGTCCCGCCGAGGCTGGGCCAGGAGCTCCA 1308
Db 165 GCCTTGGGACATCGCAGGCGGGAGAGCAGTCCCGCCGAGGCTGGGCCAGGAGCTCCA 106
QY 1309 GGAAGGGCACTGAGCGCTGTGGCGGAGGCTCGGACATCCGAGGCCAGGAAAGT 1368
Db 105 GGAAGGGCACTGAGCGCTGTGGCGGAGGCTCGGACATCCGAGGCCAGGAAAGT 46
QY 1369 CTCCTGGGGCGATCTGTAAATAAACCTTTTCTTTTCTTTTCTTTT 1413
Db 45 CTCCTGGGGCGATCTGTAAATAAACCTTTTCTTTTCTTTTCTTTT 1

Qy 916 GGGGTCTTGGCCTCCATCTCTAGCGTGGCGCTCTGCCCGCTCCGCTAGGCTCC 975
Db 181 GGGGTCTTGGCCTCCATCTCTAGCGTGGCGCTCTGCCCGCTCCGCTAGGCTCC 240
Qy 976 TCCGGCTCACCACCTAGTACGGCGCGCTTCTGCGTACGCTGCGCAACCGGCGCTCG 1035
Db 241 TCCGGCTCACCACCTAGTACGGCGCGCTTCTGCGTACGCTGCGCAACCGGCGCTCG 300
Qy 1036 TGCCTCTTCTCGGAGGCGCGCTGCTGAGTCTCCAGTATGTTGCGCCAGCGCTCTTCGC 1095
Db 301 TGCCTCTTCTCGGAGGCGCGCTGCTGAGTCTCCAGTATGTTGCGCCAGCGCTCTTCGC 360
Qy 1096 ACCCTCTGAGCAACGCGCAAGGACTG 1124
Db 361 ACCCTCTGAGCAACGCGCAAGGACTG 389
RESULT 9
AI660493/c
LOCUS
DEFINITION
we67h02.x1 Soares Dieckgraefe colon NHCD Homo sapiens cDNA clone
IMAGE:2346195 3' Similar to contains TAR1.t3 TAR1 repetitive
element ; mRNA sequence.
ACCESSION
AI660493
VERSION
AI660493.1 GI:4764063
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 538)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 676 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 463.
Location/Qualifiers
1. .538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2346195"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares Dieckgraefe colon NHCD"
/note="Organ: colon; Vector: pF773D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCGCTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pF773 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraefe (Washington University,
dieck@im.wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo."
ORIGIN
Query Match 26.9%; Score 388; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 7.1e-68;
Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1027 GGGGTCTTGGCCTCCATCTCTAGCGTGGCGCTCTGCCCGCTCCGCTAGGCTCC 1086

RESULT 8
AI791844
LOCUS
DEFINITION
nk08a11.y5 NCI CGAP Co2 Homo sapiens cDNA clone IMAGE:1012892 5'
similar to contains Alu repetitive element; mRNA sequence.
ACCESSION
AI791844
VERSION
AI791844.1 GI:5339486
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 505)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: nk08a11.s1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: see original entry for original citation
information
This 5' resequenced clone has no previous 5' data to verify this
new read against
Putative full length read
The vector to vector length is 671
Insert Length: 741 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
1. .505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1012892"
/tissue_type="tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP Co2"
/note="Organ: colon; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence:
5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."
ORIGIN
Query Match 27.0%; Score 389; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 4.8e-68;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 736 AAGTTCACACCGAGTAGCGCTTGGCGCTGTACCACTACCACTGGCGGACTAC 795
Db 1 AAGTTCACACCGAGTAGCGCTTGGCGCTGTACCACTACCACTGGCGGACTAC 60
Qy 796 GCCTCGGCCACGCTATGGTGGCGTTCTGCTCTTGGTCTCTCCACGCTGCTCTCC 855
Db 61 GCCTCGGCCACGCTATGGTGGCGTTCTGCTCTTGGTCTCTCCACGCTGCTCTCC 120
Qy 856 ACGCCGCGCCGCTCTACGAGGCGCTTGGCACTGCTGACACCGGAGCCTTCGGCTCTTC 915
Db 121 ACGCCGCGCCGCTCTACGAGGCGCTTGGCACTGCTGACACCGGAGCCTTCGGCTCTTC 180

/clone="IMAGE:1873190"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Co8"
/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 26.4%; Score 381; DB 9; Length 381;
Best Local Similarity 100.0%; Pred. No. 2.4e-66;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1033 CTGTGCTCTTCTCTCGAGGGGCGCTGTGAGTCTCCAGTATGTTCGGCCCGAGGCTCTT 1092
DB 381 CTGTGCTCTTCTCTCGAGGGGCGCTGTGAGTCTCCAGTATGTTCGGCCCGAGGCTCTT 322
QY 1093 CGCACCTTCTTGACCAAGCGCCAGGACTGCAGCCAGGAGAGGGGGCTCACCTCTT 1152
DB 321 CGCACCTTCTTGACCAAGCGCCAGGACTGCAGCCAGGAGAGGGGGCTCACCTCTT 262
QY 1153 ATCTCTGGCGACCCCACTGCACAGCAGGCGCTCTCCAGACTTAAATGTATCACT 1212
DB 261 ATCTCTGGCGACCCCACTGCACAGCAGGCGCTCTCCAGACTTAAATGTATCACT 202
QY 1213 AACCTGTGAGGGGACCCCAATCTGACTCCTTCCCGGCTTGGACATCGAGCCGGGA 1272
DB 201 AACCTGTGAGGGGACCCCAATCTGACTCCTTCCCGGCTTGGACATCGAGCCGGGA 142
QY 1273 AGCATGTCCTGGAGGCTGGGCGGAGAGCTCCAGGAAGGCACTGAGCGCTGCTGGC 1332
DB 141 AGCATGTCCTGGAGGCTGGGCGGAGAGCTCCAGGAAGGCACTGAGCGCTGCTGGC 82
QY 1333 GCGAGGCTCGGACATCCGAGGCGCAGGGAAGTCTCTGGGCGCATCTGTAAATAA 1392
DB 81 GCGAGGCTCGGACATCCGAGGCGCAGGGAAGTCTCTGGGCGCATCTGTAAATAA 22
QY 1393 CCTTTTCTTTTGTGTTTTT 1413
DB 21 CCTTTTCTTTTGTGTTTTT 1

RESULT 12
CB854121/c
LOCUS
DEFINITION
UI-CF-DUI-aal-1-16-0-UI.81 UI-CF-DUI Homo sapiens cDNA clone
UI-CF-DUI-aal-1-16-0-UI 3', mRNA sequence.
CB854121
CB854121.1 GI:30044498
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 573)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
PUBMED
8889548
COMMENT
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.regen.com).
The following repetitive elements were found in this cDNA
sequence: S33-S73, >ALU (matched complement)
Seq primer: M13 FORWARD
POLYA=No.

FEATURES

source

Location/Qualifiers
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-aal-1-16-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DUI"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT73-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG_SEQ=None found"

ORIGIN

Query Match 26.0%; Score 375; DB 14; Length 573;
Best Local Similarity 100.0%; Pred. No. 2.4e-65;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1044 CCTCGAGGGCGGTGTGAGTCTCCAGTATGTTCGGCCCGAGGCTCTTCGACCCCTTCT 1103
DB 482 CCTCGAGGGCGGTGTGAGTCTCCAGTATGTTCGGCCCGAGGCTCTTCGACCCCTTCT 423
QY 1104 GGACCAAGCGCCAAAGGACTGCAGCCAGGAGAGGGGGTCACTCTTATCTCTGGCGA 1163
DB 422 GGACCAAGCGCCAAAGGACTGCAGCCAGGAGAGGGGGTCACTCTTATCTCTGGCGA 363
QY 1164 CCCACTGCAAGCAGGCGGCTCTCCAGACTTAAATGTATCACCCTAAGCTGTGAGG 1223
DB 362 CCCACTGCAAGCAGGCGGCTCTCCAGACTTAAATGTATCACCCTAAGCTGTGAGG 303
QY 1224 GGGACCCCAATCTGGACTCCTTCCCGCTTGGACATCGCAGGCGGGAGCAGTGCCTG 1283
DB 302 GGGACCCCAATCTGGACTCCTTCCCGCTTGGACATCGCAGGCGGGAGCAGTGCCTG 243
QY 1284 CCAGGCTTGGCCAGGAGAGTCCAGGAAGGCACTGAGCGCTGCGCGAGGCGCTCG 1343
DB 242 CCAGGCTTGGCCAGGAGAGTCCAGGAAGGCACTGAGCGCTGCGCGAGGCGCTCG 183
QY 1344 GACATCGCAGGCGCAGGAAAGTCTCTCGGGCGGATCTGTAAATAAACCTTTTCT 1403
DB 182 GACATCGCAGGCGCAGGAAAGTCTCTCGGGCGGATCTGTAAATAAACCTTTTCT 123
QY 1404 TTTGTTTTTAAAAA 1418
DB 122 TTTGTTTTTAAAAA 108
RESULT 13
AI346155/c
LOCUS
AI346155 405 bp mRNA linear EST 02-FEB-1999

DEFINITION qp43f12.x1 NCI CGAP Co8 Homo sapiens cDNA clone IMAGE:1925807 3' similar to contains_TAR1.t3 TAR1 repetitive element ;, mRNA

AI346155
VERSION AI346155.1 GI:4083361
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 405)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1488 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 375.
Location/Qualifiers
FEATURES
source
1..405
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1925807"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Co8"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 26.0%; Score 374; DB 9; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.4e-65;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1009 TGGGTACAGCTGGCAACCGGCTCTGTGCCTTCTTCTCGAGGGGCGGTGGTGTCTC 1068
Db 405 TGGGTACAGCTGGCAACCGGCTCTGTGCCTTCTTCTCGAGGGGCGGTGGTGTCTC 346

QY 1069 CAGTATGTTGGGCCAGCGCTCTTGCACCCCTTCTGGACCAAGCGCCAGGACTGCAGC 1128
Db 345 CAGTATGTTGGGCCAGCGCTCTTGCACCCCTTCTGGACCAAGCGCCAGGACTGCAGC 286

QY 1129 CAGGAGAGAGGGGCTCACCTCTTATCTCTGGGACCCACTGCACAGCAGGCGGTCTC 1188
Db 285 CAGGAGAGAGGGGCTCACCTCTTATCTCTGGGACCCACTGCACAGCAGGCGGTCTC 226

QY 1189 CCAGACTTAAATGTATCACCACCTAACCCTGTGAGGGGGACCCAACTGTGGACTCTTCCOC 1248
Db 225 CCAGACTTAAATGTATCACCACCTAACCCTGTGAGGGGGACCCAACTGTGGACTCTTCCOC 166

QY 1249 GCCTTGGGACATCGCAGGCGGGAGAGCAGTGCCCGCCAGGCGCTGGGCCAGGAGCTCCA 1308
Db 165 GCCTTGGGACATCGCAGGCGGGAGAGCAGTGCCCGCCAGGCGCTGGGCCAGGAGCTCCA 106

QY 1309 GGAAGGGCACTGAGCGCTGTGGCGGAGGCTCTGGACATCCGAGGCACCCAGGAAAGT 1368

Db 105 GGAAGGGCACTGAGCGCTGTGGCGGAGGCGCTCGGACATCCGAGGCCACGAGAAAGT 46

QY 1369 CTCCTGGGGCGATC 1382
Db 45 CTCCTGGGGCGATC 32

RESULT 14
AI660560/c
LOCUS AI660560
DEFINITION we68b08.x1 Soares Dieckgraefe_colon_NHCD Homo sapiens cDNA clone IMAGE:2346231 3' similar to contains_TAR1.t3 TAR1 repetitive element ;, mRNA sequence.
ACCESSION AI660560
VERSION AI660560.1 GI:4764130
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 486)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-re@mail.nih.gov
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1433 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 447.
Location/Qualifiers
FEATURES
source
1..486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2346231"
/tissue_type="colonic mucosa from 3 patients with Crohn's disease"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares Dieckgraefe colon_NHCD"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer {5', TGTTACCAATCTGAAGTGGAGCGCGCGCTCTTTTCTTTTCTTTT 3'}, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieck@wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 25.6%; Score 369; DB 9; Length 486;
Best Local Similarity 99.8%; Pred. No. 4.4e-64;
Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1009 TGGGTACAGCTGGCAACCGGCTCTTGTGSCCTTCTTCTCGAGGGGCGGTGGTGTCTC 1068
Db 420 TGGGTACAGCTGGCAACCGGCTCTTGTGSCCTTCTTCTCGAGGGGCGGTGGTGTCTC 361

QY 1069 CAGTATGTTGGGCCAGCGCTCTTGCACCCCTTCTGGACCAAGCGCCAGGACTGCAGC 1128
Db 360 CAGTATGTTGGGCCAGCGCTCTTGCACCCCTTCTGGACCAAGCGCCAGGACTGCAGC 301

QY 1129 CAGGAGAGAGGGGCTCACCTCTTATCTCTGGCGGACCCACTGCACAGGCGGTCTC 1188
Db 300 CAGGAGAGAGGGGCTCACCTCTTATCTCTGGCGGACCCACTGCACAGGCGGTCTC 241

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 13:19:30 ; Search time 653 Seconds
(without alignments)
10067.131 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 2995936 seqs, 2280998010 residues

Word size : 0
Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1413	98.1	1421	15	US-10-187-657-2	Sequence 2, Appli
2	1404	97.4	1420	15	US-10-187-657-4	Sequence 4, Appli
3	1285	89.2	1474	16	US-10-264-237-1097	Sequence 1097, Ap
C	392	27.2	2684	16	US-10-094-749-1195	Sequence 1195, Ap
	5	313	21.7	522	15	US-10-187-657-5
6	187	13.0	232	9	US-09-783-590-9492	Sequence 9492, Ap
7	136	9.4	346	15	US-10-187-657-3	Sequence 3, Appli
8	119	8.3	450	15	US-10-187-657-7	Sequence 7, Appli
9	107	7.4	439	13	US-10-027-632-91970	Sequence 91970, A
10	107	7.4	439	16	US-10-027-632-91970	Sequence 91970, A
11	59	4.1	512	15	US-10-106-698-3277	Sequence 3277, Ap
12	56	3.9	506	15	US-10-187-657-6	Sequence 6, Appli
13	38	2.6	999	12	US-09-876-143-1614	Sequence 1614, Ap
14	38	2.6	1398	13	US-10-424-599-64755	Sequence 64755, A

15	38	2.6	2165	15	US-10-116-255-18	Sequence 18, Appl
16	38	2.6	3095	13	US-10-424-599-75074	Sequence 75074, A
17	38	2.6	5642	15	US-10-311-455-842	Sequence 842, App
18	38	2.6	10480	15	US-10-311-455-2174	Sequence 2174, Ap
19	37	2.6	554	13	US-10-424-599-41481	Sequence 41481, A
20	37	2.6	2484	13	US-10-424-599-8115	Sequence 8115, Ap
C	21	37	3673778	15	US-10-312-841-1	Sequence 1, Appli
	22	36	896	15	US-10-106-698-234	Sequence 234, App
23	35	2.4	290	13	US-10-085-783A-52594	Sequence 52594, A
24	35	2.4	290	16	US-10-242-535A-52594	Sequence 52594, A
C	25	35	302	13	US-10-424-599-48927	Sequence 48927, A
	26	35	379	13	US-10-424-599-43341	Sequence 43341, A
27	35	2.4	379	13	US-10-424-599-133773	Sequence 133773, A
28	35	2.4	416	13	US-10-424-599-50281	Sequence 50281, A
29	35	2.4	449	13	US-10-424-599-1341	Sequence 1341, Ap
30	35	2.4	450	15	US-10-198-846-13053	Sequence 13053, A
31	35	2.4	450	16	US-10-131-827-8647	Sequence 8647, Ap
32	35	2.4	464	16	US-10-131-827-8179	Sequence 8179, Ap
33	35	2.4	478	13	US-10-424-599-22997	Sequence 22997, A
34	35	2.4	481	13	US-10-424-599-120735	Sequence 120735, A
35	35	2.4	482	13	US-10-424-599-96207	Sequence 96207, A
36	35	2.4	494	13	US-10-424-599-17387	Sequence 17387, A
37	35	2.4	498	13	US-10-424-599-17873	Sequence 17873, A
38	35	2.4	509	13	US-10-424-599-107007	Sequence 107007, A
39	35	2.4	529	15	US-10-029-386-3130	Sequence 3130, Ap
40	35	2.4	621	13	US-10-424-599-10941	Sequence 10941, A
C	41	35	658	13	US-10-424-599-122157	Sequence 122157, A
	42	35	1167	13	US-10-027-632-260991	Sequence 260991, A
43	35	2.4	1167	16	US-10-027-632-260991	Sequence 260991, A
44	35	2.4	1639	9	US-09-742-954-5	Sequence 5, Appli
45	35	2.4	1758	13	US-10-424-599-91053	Sequence 91053, A

ALIGNMENTS

RESULT 1
US-10-187-657-2
; Sequence 2, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Maria R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 4901066CB1
US-10-187-657-2

Query Match 98.1%; Score 1413; DB 15; Length 1421;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAAGTAACGGCTACAGACAGTACAGAAATAGTTCTGCTCGCGGCTAGAAAACCTCTGTG 60
DB 9 AAAGTAACGGCTACAGACAGTACAGAAATAGTTCTGCTCGCGGCTAGAAAACCTCTGTG 68
QY 61 GTACCAACCCGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGAGAGGTGCA 120

Db	69	GTACCAACCCACAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAAACGAGAGGTGCA	128
QY	121	GGACTCAGACTTCAACGAGCCACTCGGTCCAGCCCTTGTAAGAAAGAGACGCCAAGGAC	180
Db	129	GGACTCAGACTTCAACGAGCCACTCGGTCCAGCCCTTGTAAGAAAGAGACGCCAAGGAC	188
QY	181	GCGCTCTCCGCGTCCAGGAGCCCAAGCTTGCTGGCTTGCCTGCCCGCTGCGTGCAGC	240
Db	189	GCGCTCTCCGCGTCCAGGAGCCCAAGCTTGCTGGCTTGCCTGCCCGCTGCGTGCAGC	248
QY	241	ACTCGCCGCGTGCAGCATGACCTGTGGAACGGCGTACTGCCTTTTATACCCCCAGCCC	300
Db	249	ACTCGCCGCGTGCAGCATGACCTGTGGAACGGCGTACTGCCTTTTATACCCCCAGCCC	308
QY	301	CGGCATGCCGAGGCTTCAGCGTCCACTGCTCATCGTTATTCTAGTGTTTTGGCTCTA	360
Db	309	CGGCATGCCGAGGCTTCAGCGTCCACTGCTCATCGTTATTCTAGTGTTTTGGCTCTA	368
QY	361	GCAGCAAGCTTCTGCTCATCTTGCCGGGGAFCGGGCACTCGCGCTGCTTTTGGTTG	420
Db	369	GCAGCAAGCTTCTGCTCATCTTGCCGGGGAFCGGGCACTCGCGCTGCTTTTGGTTG	428
QY	421	GTGAGAGTTCTTCTCAGTCTGTTTATAGGCGAGAAATTGTGGCTGTGCACTTCAGTGCA	480
Db	429	GTGAGAGTTCTTCTCAGTCTGTTTATAGGCGAGAAATTGTGGCTGTGCACTTCAGTGCA	488
QY	481	GAATGGTTCGCGGTACAGTGAACACCAACACATCCTACAAAGCCTTCAGCGCAGCGGC	540
Db	489	GAATGGTTCGCGGTACAGTGAACACCAACACATCCTACAAAGCCTTCAGCGCAGCGGC	548
QY	541	GTTACAGCCCGTGTGCGTCTGCTCGTGGGCTGAGGGCAATTAATATTACACTCACAGG	600
Db	549	GTTACAGCCCGTGTGCGTCTGCTCGTGGGCTGAGGGCAATTAATATTACACTCACAGG	608
QY	601	ACCCAGTGCACTCAGCTGAACAGAGCAATTTGACTACAAAGAGGGGCTGCCGACCCAGTG	660
Db	609	ACCCAGTGCACTCAGCTGAACAGAGCAATTTGACTACAAAGAGGGGCTGCCGAGCTG	668
QY	661	AAAGAGAAATTACGCGCGGAGTACGCGAACCGACTGGAGAGGGGCTGCCGACCCAGTG	720
Db	669	AAAGAGAAATTACGCGCGGAGTACGCGAACCGACTGGAGAGGGGCTGCCGACCCAGTG	728
QY	721	CTCTACCTGGCGGAGAAATTACACACGAGTAGCCCTTGGCGGCTGTACCAACAGTACCAC	780
Db	729	CTCTACCTGGCGGAGAAATTACACACGAGTAGCCCTTGGCGGCTGTACCAACAGTACCAC	788
QY	781	CTGGCGGACACTAGCCCTCGGCCACGCTATGGGTGGCGTTCTGCTCTGGCTCCTCTCC	840
Db	789	CTGGCGGACACTAGCCCTCGGCCACGCTATGGGTGGCGTTCTGCTCTGGCTCCTCTCC	848
QY	841	AACGTGCTGCTCTCCACGCGGCCCGCTCTACGAGGCGCTTGCACTGCTGACCAACCGGA	900
Db	849	AACGTGCTGCTCTCCACGCGGCCCGCTCTACGAGGCGCTTGCACTGCTGACCAACCGGA	908
QY	901	GCTTCGCGCTCTTGGGGTCTTGGCTTGCCCTCCATCTCTAGCGTGGCGCTCTGCCCG	960
Db	909	GCTTCGCGCTCTTGGGGTCTTGGCTTGCCCTCCATCTCTAGCGTGGCGCTCTGCCCG	968
QY	961	CTCGGCTTAGGCTCCTCGCGCTCACCACTCAGTACGGCGCGCTTCTGGGTCAACGCTG	1020
Db	969	CTCGGCTTAGGCTCCTCGCGCTCACCACTCAGTACGGCGCGCTTCTGGGTCAACGCTG	1028
QY	1021	GCAACCGGCTCCTGTGCCTCTTCTCGGAGGGCGGTGGTGAGTCTCCAGTATGTTCCG	1080
Db	1029	GCAACCGGCTCCTGTGCCTCTTCTCGGAGGGCGGTGGTGAGTCTCCAGTATGTTCCG	1088
QY	1081	CCACAGCGCTCTTCCGACCCCTTCTGGACCAAAGCGCAAGGACTGCAGCCAGAGAGGG	1140
Db	1089	CCACAGCGCTCTTCCGACCCCTTCTGGACCAAAGCGCAAGGACTGCAGCCAGAGAGGG	1148
QY	1141	GGCTCACCTCTTATCCTCGGCGACCCACTGCACAAGCAGGCCGCTCTCCAGACTTAAA	1200
Db	1149	GGCTCACCTCTTATCCTCGGCGACCCACTGCACAAGCAGGCCGCTCTCCAGACTTAAA	1208

RESULT 2

RESUL 2
US-10-187-657-4

US-10-187-631-4
: Sequence 4, Application US/10187657

Publication No. US20030068311A1

GENERAL INFORMATION:

APPLICANT: Lasek, Amy K.W.

APPLICANT: Dasek, Amy K.W.
APPLICANT: Baughn, Mariah B.

APPLICANT: Azimzai, Yalda

APPLICANT: AZIMZAL, IAIQA
TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER

FILE OF INVENTION: IRANSMER
FILE REFERENCE: PV-0009 CIP

FILE REFERENCE: PV-0003 CIF
CURRENT APPLICATION NUMBER: US/10/187.657

; CURRENT APPLICATION NUMBER: US11
 : CURRENT FILING DATE: 2002-07-01

; CURRENT FILING DATE: 2002-07-01
 ;
 : PRIOR APPLICATION NUMBER: PCT/US00/07817

; PRIOR APPLICATION NUMBER: FCI-
 : PRIOR FILING DATE: 2000 03-22

;; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139-565

PRIOR FILING DATE: 1999-06-16
PRIOR APPLICATION NUMBER: 607116

; PRIOR FILING DATE: 1999-
 : NUMBER OF SEQ IN NOS: 9

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: I

; SEQ ID NO 4
1 ENCTY 1478; LENGTH: 14
TYPE: DNA

TYPE: DNA

; ORGANISM

;
FEATURE:
11/11/11 11:11

```
; NAME/KEY: misc_feature  
; COUNTER INFORMATION: TEST#0 ID NO: JSC00030068311A1 3221667CA2
```

OTHER INFORMATION

Query Match 97.48; Score 1404; DB 15; Length 1420;

Query Match	57.14%	Score 1.00
Best Local Similarity	100.0%	Pred. No. 0;

Best Local Similarity	100.00%	Recall	100.00%
Matches 1404:	Conservative	0;	Mismatches
		0;	Indels
		0;	Gaps

1 AAAGTAACGGCTACAGACAGTGAGAAATAGTTTCGGCTCGCCGGCTAGAAAACTCTGTGG 60

100

61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGGAGAGGTGCA 120

01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

121 GGACTCAGACTTCACCCAGCCCCACTCGGTCCAGCCTTGTACGCAAGAGACGCCAAGGAC 180

[illegible]

181 GCGCTCTCCCGCGTCCAGGCAGCCCCAGCTTGCTGGCTTGCCCGCCGCGCTGCGTGCAGC 240

[illegible]

241 ACTCGGCCGGCTGCAGCATGACCCCTGTGGAACGGGTA CTGCCCTTTTACCCCCAGCCC 300

[illegible]

301 CGGCATGCCCGAGGGTTACGGTTCCACCTGGTCATCGTTATTCCTAGTGTCTTTTGGCTCTA 360

301 CGGATGCCGAGGCTTCAGGTTCCACAGCTACACGTTATTCATGAGTATTCGCTCTC
300


```
QY 721 CTCTACCTGGCGGAGAAATTCAACCGAGTAGCCCTTGGGCTGTACACAGTACAC 780
Db 747 CTCTACCTGGCGGAGAAATTCAACCGAGTAGCCCTTGGGCTGTACACAGTACAC 806
QY 781 CTGGCGGAGACTACGCTTGGCCACGCTATGGGTGGGCTTCTGCTTCTGGCTCTCTCC 840
Db 807 CTGGCGGAGACTACGCTTGGCCACGCTATGGGTGGGCTTCTGCTTCTGGCTCTCTCC 866
QY 841 AACGTGCTGCTCTCCACGCGCGCCGCTCTACGAGGCGCTGGCACTGCTGACCAACGGA 900
Db 867 AACGTGCTGCTCTCCACGCGCGCCGCTCTACGAGGCGCTGGCACTGCTGACCAACGGA 926
QY 901 GCCTTGGCGCTCTTGGGCTCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 960
Db 927 GCCTTGGCGCTCTTGGGCTCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 986
QY 961 CTCCGCTAGGCTCTCCGCGCTTCCACCTACGACTACGCGCGGCTTCTGGCTCAAGCTG 1020
Db 987 CTCCGCTAGGCTCTCCGCGCTTCCACCTACGACTACGCGCGGCTTCTGGCTCAAGCTG 1046
QY 1021 GCAACCGGCGCTCTGCTCTCTTCTCGGAGGCGCGCTGGTGAGTCTCCAGTATGTTCCG 1080
Db 1047 GCAACCGGCGCTCTGCTCTCTTCTCGGAGGCGCGCTGGTGAGTCTCCAGTATGTTCCG 1106
QY 1081 CCAGCGCTCTTCCGACCCCTTCTGGACCAAGCGCCCAAGGACTGCGAGGAGAGGG 1140
Db 1107 CCAGCGCTCTTCCGACCCCTTCTGGACCAAGCGCCCAAGGACTGCGAGGAGAGGG 1166
QY 1141 GGCTCACCTCTTATCTCGGCGACCCACTGCAAGCAGGCGCGCTCTCCAGACTTAAAA 1200
Db 1167 GGCTCACCTCTTATCTCGGCGACCCACTGCAAGCAGGCGCGCTCTCCAGACTTAAAA 1226
QY 1201 TGATACCACTAACCTGTGAGGGGACCCCAATCTGGAATCTCCCTCCCGCTTGGGACAT 1260
Db 1227 TGATACCACTAACCTGTGAGGGGACCCCAATCTGGAATCTCCCTCCCGCTTGGGACAT 1286
QY 1261 CGAGGCGGGAAGCAGTCCCGCAGGCGCTGGGCGAGGAGCTCCAGGAGGCGCACTG 1320
Db 1287 CGAGGCGGGAAGCAGTCCCGCAGGCGCTGGGCGAGGAGCTCCAGGAGGCGCACTG 1346
QY 1321 AGCGTGTGGCGGAGGCGCTCGGACATCCGAGGCGCAGGGAAGTCTCTGGGGCGGA 1380
Db 1347 AGCGTGTGGCGGAGGCGCTCGGACATCCGAGGCGCAGGGAAGTCTCTGGGGCGGA 1406
QY 1381 TCTGTAATAAACCTTTTCTTTCTTTTAAAAA 1438
Db 1407 TCTGTAATAAACCTTTTCTTTCTTTTAAAAA 1464
```

RESULT 4
US-10-094-749-1195/c
; Sequence 1195, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO

```
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1195  
; LENGTH: 2684  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-094-749-1195  
  
Query Match 27.2%; Score 392; DB 16; Length 2684;  
Best Local Similarity 100.0%; Pred. No. 2.6e-186;  
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1027 GGGTCTCTGTGCTCTTCTCTCGGAGGGCCGCTGGTGAGTCTCCAGTATGTTCCGGCCCGC 1086  
Db 2343 GGGTCTCTGTGCTCTTCTCTCGGAGGGCCGCTGGTGAGTCTCCAGTATGTTCCGGCCCGC 2284  
QY 1087 GCTCTTCGACCCCTTCTGGACCAAGCGCCCAAGGACTGCGAGGAGAGAGGGGGCTCA 1146  
Db 2283 GCTCTTCGACCCCTTCTGGACCAAGCGCCCAAGGACTGCGAGGAGAGAGGGGGCTCA 2224  
QY 1147 CCTCTTATCTCTCGGCGACCCACTGCAAGCAGGCGCGCTCTCCAGACTTAAATGTATC 1206  
Db 2223 CCTCTTATCTCTCGGCGACCCACTGCAAGCAGGCGCGCTCTCCAGACTTAAATGTATC 2164  
QY 1207 ACCACTAACCTGTGAGGGGACCCCAATCTGGAATCTCCCTCCCGCTTGGGACATCGCAGG 1266  
Db 2163 ACCACTAACCTGTGAGGGGACCCCAATCTGGAATCTCCCTCCCGCTTGGGACATCGCAGG 2104  
QY 1267 CCGGGAAGCAGTCCCGCAGGCGCTGGGCGAGGAGCTCCAGGAAGGCGACTGAGCGCT 1326  
Db 2103 CCGGGAAGCAGTCCCGCAGGCGCTGGGCGAGGAGCTCCAGGAAGGCGACTGAGCGCT 2044  
QY 1327 GCTGGCGGAGGCGCTCGGACATCCGAGGCGCAGGGAAGTCTCTCTGGGGCGATCTGTA 1386  
Db 2043 GCTGGCGGAGGCGCTCGGACATCCGAGGCGCAGGGAAGTCTCTCTGGGGCGATCTGTA 1984  
QY 1387 AATAAACCTTTTCTTTTCTTTTAAAAA 1418  
Db 1983 AATAAACCTTTTCTTTTCTTTTAAAAA 1952
```

RESULT 5
US-10-187-657-5
; Sequence 5, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzal, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. US20030068311A1 3221661R6

FEATURE: 100.0%; Score 313; DB 15; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.4e-146;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
NAME/KEY: unsure
LOCATION: 455, 480, 483
OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-5

Query Match 21.7%; Score 313; DB 15; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.4e-146;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTCGCGGCTAGAAAAACTCTGTG 60
Db 9 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTCGCGGCTAGAAAAACTCTGTG 68
Qy 61 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCAGCTTCTTAAACGGAGAGTGCA 120
Db 69 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCAGCTTCTTAAACGGAGAGTGCA 128
Qy 121 GGAATCAGACTTCAACGAGCCCACTCGGTCCAGCCCTTGTACGCAAGAGAGCCCAAGGAC 180
Db 129 GGAATCAGACTTCAACGAGCCCACTCGGTCCAGCCCTTGTACGCAAGAGAGCCCAAGGAC 188
Qy 181 GCGCTCTCCGCGTCCAGGAGCCCACTCGGTCCAGCCCTTGTACGCAAGAGAGCCCAAGGAC 240
Db 189 GCGCTCTCCGCGTCCAGGAGCCCACTCGGTCCAGCCCTTGTACGCAAGAGAGCCCAAGGAC 248
Qy 241 ACTCGGCGCGGTGCAGCATGACCCCTGTGGAACGGCGTACTGCCCTTTTACCCCGAGCCC 300
Db 249 ACTCGGCGCGGTGCAGCATGACCCCTGTGGAACGGCGTACTGCCCTTTTACCCCGAGCCC 308
Qy 301 CGGCATGCCGAG 313
Db 309 CGGCATGCCGAG 321

RESULT 6

US-09-783-590-9492
Sequence 9492, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2C1
CURRENT APPLICATION NUMBER: US/09/783,590
CURRENT FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIOR FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9492
LENGTH: 232
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (202)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (206)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (229)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9492

Query Match

13.0%; Score 187; DB 9; Length 232;

Best Local Similarity 100.0%; Pred. No. 2.7e-83;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTCGCGGCTAGAAAAACTCTGTG 60
Db 15 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTCGCGGCTAGAAAAACTCTGTG 74
Qy 61 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCAGCTTCTTAAACGGAGAGTGCA 120
Db 75 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCAGCTTCTTAAACGGAGAGTGCA 134
Qy 121 GGAATCAGACTTCAACGAGCCCACTCGGTCCAGCCCTTGTACGCAAGAGAGCCCAAGGAC 180
Db 135 GGAATCAGACTTCAACGAGCCCACTCGGTCCAGCCCTTGTACGCAAGAGAGCCCAAGGAC 194
Qy 181 GCGCTCT 187
Db 195 GCGCTCT 201

RESULT 7

US-10-187-657-3
Sequence 3, Application US/10187657
Publication No. US20030068311A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy K.W.
APPLICANT: Baughn, Mariah R.
APPLICANT: Azimzai, Valda
TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
FILE REFERENCE: PV-0009 CIP
CURRENT APPLICATION NUMBER: US/10/187,657
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: PCT/US00/07817
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/139,565
PRIOR FILING DATE: 1999-06-16
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PERL Program
SEQ ID NO 3
LENGTH: 346
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030068311A1 1752794F6
NAME/KEY: unsure
LOCATION: 91, 188, 206, 291, 337
OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-3

Query Match 9.4%; Score 136; DB 15; Length 346;
Best Local Similarity 99.5%; Pred. No. 1.1e-57;
Matches 186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 927 CTTGGCTCCATCTCTAGCGTGCCGCTCTGCGGCTCGGCTAGGCTCCTCGCGGCTCAC 986
Db 1 CTTGGCTCCATCTCTAGCGTGCCGCTCTGCGGCTCGGCTAGGCTCCTCGCGGCTCAC 60
Qy 987 CACTCAGTACGGCGCGCTTCTGGGTACCGCTGGCAACCGGCGTCTGCTCTTCT 1046
Db 61 CACTCAGTACGGCGCGCTTCTGGGTACCGCTGGCAACCGGCGTCTGCTCTTCT 120
Qy 1047 CGGAGGGCGCGTGGTGGTCTCCAGTATGTTGGCCAGCGGCTCTTTCGACCCCTTCTGGA 1106
Db 121 CGGAGGGCGCGTGGTGGTCTCCAGTATGTTGGCCAGCGGCTCTTTCGACCCCTTCTGGA 180
Qy 1107 CCAAGC 1113
Db 181 CCAAGC 187

RESULT 8

```

US-10-187-657-7
; Sequence 7, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 SEQID4510D1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 140, 165, 203, 237, 269, 289, 307, 338, 344, 347, 354, 390, 430, 445
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-7

Query Match      8.3%; Score 119; DB 15; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.6e-49;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 CTGAACGAGACCAATTGACTACAAACGAGCAGTTACCTGGGCTCTGAAGAGAGATTACGCC 675
    |||||||
Db 21 CTGAACGAGACCAATTGACTACAAACGAGCAGTTACCTGGGCTCTGAAGAGAGATTACGCC 80

QY 676 GCGGAGTACGGAACGCACTGGAGAGGGGCTGCCGACCCAGTGTCTTACCTGGCGGA 734
    |||||||
Db 81 GCGGAGTACGGAACGCACTGGAGAGGGGCTGCCGACCCAGTGTCTTACCTGGCGGA 139

RESULT 9
US-10-027-632-91970
; Sequence 91970, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91970
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-91970

Query Match      7.4%; Score 107; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.9e-43;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 ACCACCGAGCCTTCGGCTCTTCGGGGTCTTCGGCTTCGGCTTCATCTCTAGCGTCCG 951
    |||||||
Db 214 ACCACCGAGCCTTCGGCTCTTCGGGGTCTTCGGCTTCGGCTTCATCTCTAGCGTCCG 273

QY 952 CTCTGCCCGCTCCGCTAGGCTCCTCCGGCTCACCCTCAGTACGG 998
    |||||||
Db 274 CTCTGCCCGCTCCGCTAGGCTCCTCCGGCTCACCCTCAGTACGG 320

RESULT 11
US-10-106-698-3277
; Sequence 3277, Application US/10106698

```

```

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-91970

Query Match      7.4%; Score 107; DB 13; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.9e-43;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 ACCACCGAGCCTTCGGCTCTTCGGGGTCTTCGGCTTCGGCTTCATCTCTAGCGTCCG 951
    |||||||
Db 214 ACCACCGAGCCTTCGGCTCTTCGGGGTCTTCGGCTTCGGCTTCATCTCTAGCGTCCG 273

QY 952 CTCTGCCCGCTCCGCTAGGCTCCTCCGGCTCACCCTCAGTACGG 998
    |||||||
Db 274 CTCTGCCCGCTCCGCTAGGCTCCTCCGGCTCACCCTCAGTACGG 320

RESULT 10
US-10-027-632-91970
; Sequence 91970, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 10827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91970
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-91970

Query Match      7.4%; Score 107; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.9e-43;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 ACCACCGAGCCTTCGGCTCTTCGGGGTCTTCGGCTTCGGCTTCATCTCTAGCGTCCG 951
    |||||||
Db 214 ACCACCGAGCCTTCGGCTCTTCGGGGTCTTCGGCTTCGGCTTCATCTCTAGCGTCCG 273

QY 952 CTCTGCCCGCTCCGCTAGGCTCCTCCGGCTCACCCTCAGTACGG 998
    |||||||
Db 274 CTCTGCCCGCTCCGCTAGGCTCCTCCGGCTCACCCTCAGTACGG 320

RESULT 11
US-10-106-698-3277
; Sequence 3277, Application US/10106698

```

```
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 3277
LENGTH: 512
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: {392}..(392)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: {453}..(453)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: {461}..(461)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: {493}..(493)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: {499}..(499)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: {509}..(509)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-3277

Query Match 4.1%; Score 59; DB 15; Length 512;
Best Local Similarity 100.0%; Pred. No. 4.9e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 405 GCGCTGTTTGGTTGGTGAGAGTTCTTCTCAGTCTGTTTCATAGGCGCAGAAATTGTGG 463
Db 220 GCGCTGTTTGGTTGGTGAGAGTTCTTCTCAGTCTGTTTCATAGGCGCAGAAATTGTGG 278

RESULT 12
US-10-187-657-6
Sequence 6, Application US/10187657
Publication No. US20030068311A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy K.W.
APPLICANT: Baughn, Mariah R.
APPLICANT: Azimzai, Yalda
TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
FILE REFERENCE: PV-0009 CIP
CURRENT APPLICATION NUMBER: US/10/187,657
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: PCT/US00/07817
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/139,565
PRIOR FILING DATE: 1999-06-16
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PERL Program
SEQ ID NO 6
LENGTH: 506
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
```

```
OTHER INFORMATION: Incyte ID No. US20030068311A1 SEQ#03652D1
FEATURE:
NAME/KEY: unsure
LOCATION: 87, 89, 93, 95, 124, 178, 227, 253, 262, 279, 285, 287, 293-294, 298,
LOCATION: 300, 303, 312, 319-321, 323, 325, 327, 332-333, 336, 343-344, 349, 377, 3;
LOCATION: 405, 409-410, 412, 415, 420, 423, 429, 432, 436-437, 444-445, 448, 452, 4;
LOCATION: 466, 469-471, 479, 483, 485, 492, 494, 499, 501, 504
OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-6

Query Match 3.9%; Score 56; DB 15; Length 506;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 479 CAGAATGGTTCGTGGGTACAGTGAACACCAACACATCCTACAAAGCCTTCAGCGCA 534
Db 31 CAGAATGGTTCGTGGGTACAGTGAACACCAACACATCCTACAAAGCCTTCAGCGCA 86

RESULT 13
US-09-876-143-1614
Sequence 1614, Application US/09876143
Publication No. US20040081958A1
GENERAL INFORMATION:
APPLICANT: Infigen Inc.
APPLICANT: EILERTSEN, KENNETH J.
APPLICANT: PFISTER-GENSKOW, MARTHA
APPLICANT: CHILDS, LYNETTE
APPLICANT: FORSYTHE, TODD
APPLICANT: BISHOP, MICHAEL D.
TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
TITLE OF INVENTION: CELLULAR REPROGRAMMING
FILE REFERENCE: 028040-0202
CURRENT APPLICATION NUMBER: US/09/876,143
CURRENT FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: 60/209,874
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 1744
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1614
LENGTH: 999
TYPE: DNA
ORGANISM: Bovine
FEATURE:
NAME/KEY: misc feature
LOCATION: {1}..(999)
OTHER INFORMATION: n is a, c, g, or t
US-09-876-143-1614

Query Match 2.6%; Score 38; DB 12; Length 999;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1404 TTTGTTTTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 1441
Db 729 TTTGTTTTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 766

RESULT 14
US-10-424-599-64755
Sequence 64755, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: KOVALIC David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
```

```

; SEQ ID NO 64755
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29487C.1
US-10-424-599-64755

```

```

Query Match      2.6%; Score 38; DB 13; Length 1398;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1404 TTGTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1441
    |||||||
Db 1325 TTGTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1362

```

```

RESULT 15
US-10-116-255-18
; Sequence 18, Application US/10116255
; Publication No. US20030036646A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Uncoupling Protein Polynucleotides, Polypeptides, and
; FILE REFERENCE: PT009P1
; CURRENT APPLICATION NUMBER: US/10/116,255
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 09/685,897
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: PCT/US00/09534
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,701
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/142,821
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/149,448
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/164,751
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 2165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-255-18

```

```

Query Match      2.6%; Score 38; DB 15; Length 2165;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1404 TTGTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1441
    |||||||
Db 2128 TTGTTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2165

```

Search completed: June 3, 2004, 16:28:19
Job time : 659 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 08:19:31 ; Search time 637 Seconds
(without alignments)
9610.131 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1441	100.0	1441	3	AAA94623 Human CAS
2	1436	99.7	1491	9	ADD19228 Human cDN
3	1435.6	99.6	1474	6	ABL90535 Human pol
4	1413	98.1	1421	3	AAA96505 CDNA enco
5	1392.8	96.7	1460	3	AAC95564 Human sec
6	1392.8	96.7	1460	7	ABZ67235 Human sec
7	1392.8	96.7	1460	7	ABZ73640 Secreted
8	1392.8	96.7	1460	9	ADC20289 Human sec
9	465.4	32.3	498	3	AAA94624 Human CAS
10	405.4	28.1	5033	7	ABZ68115 Human sec
11	405.4	28.1	5033	7	ABZ74587 Secreted
12	405.4	28.1	5033	9	ADC21005 Human sec
13	405	28.1	580	9	ADD19187 Human cDN
14	401.8	27.9	406	5	AAS66563 DNA enco
15	396	27.5	2684	7	ADA53627 Human cod
16	356.4	24.7	1636	6	ABN59923 Novel hum
17	354.8	24.6	1029	5	AAP45134 Human INT
18	354.8	24.6	1029	7	ACD66787 Secreted
19	354.8	24.6	2133	5	AAP45133 Human INT
20	354.8	24.6	2133	7	ABX94107 CDNA enco
21	354.8	24.6	2133	7	ACD66786 Secreted
22	245.8	17.1	2684	7	ADA53627 Human cod
23	161.8	11.2	1338	5	AAS67707 DNA enco

24	65.6	4.6	1602	4	ABL21213	Ab121213 Drosophil
25	65.6	4.6	1808	4	ABL07775	Ab107775 Drosophil
26	64	4.4	512	4	AAH36185	Aah36185 Human col
27	60.6	4.2	2000	7	ADA71938	Ada71938 Rice gene
28	52.6	3.7	134	7	ABZ68116	Abz68116 Human sec
29	52.6	3.7	134	7	ABZ74588	Abz74588 Secreted
30	52.6	3.7	134	9	ADC21006	Adc21006 Human sec
31	52	3.6	2000	7	ADA71938	Ada71938 Rice gene
32	49	3.4	496	5	ABV56664	Abv56664 Human pro
33	48.6	3.4	1400	4	AAS01017	Aas01017 Maize dis
34	47.8	3.3	396	3	AAZ34971	Aaz34971 Wheat bet
35	47.6	3.3	277	7	ABX47508	Abx47508 Bovine ES
36	47.4	3.3	915	3	AAC60068	Aac60068 Human sec
37	46.8	3.2	656	4	AAL23653	Aal23653 Human bre
38	46.8	3.2	2484	5	AAS78178	Aas78178 DNA enco
39	46.8	3.2	2589	7	ABZ75977	Abz75977 Human G p
40	46.8	3.2	14055	7	AAL61170	Aal61170 Actinosyn
41	46.8	3.2	82746	7	AAL61224	Aal61224 Actinosyn
42	46.4	3.2	2206	4	ABK43595	Abk43595 DNA enco
43	46.2	3.2	441	3	AAC69600	Aac69600 Human sec
44	46.2	3.2	531	8	ADA49299	Ada49299 Maize gen
45	46.2	3.2	609	8	ADA49313	Ada49313 Maize gen

ALIGNMENTS

RESULT 1
AAA94623
ID AAA94623 standard; DNA; 1441 BP.

AC AAA94623;

DT 11-JAN-2001 (first entry)

DE Human CASB618 coding sequence.

XX Epitope; human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian; colon; autoimmune disease; HLA_A0201; ss.

OS Homo sapiens.

Key Location/Qualifiers
CDS 259..1221
/*tag= a
/*product= "Human CASB618"

PN WO200053748-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-EP002048.

XX 11-MAR-1999; 99GB-00005607.

XX 01-SEP-1999; 99GB-00020590.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Bruck CEM, Cassart J, Coche T, Vinals Y De BassolsC;

XX WPI; 2000-572268/53.

XX P-PSDB; AAB26325.

XX New human CASB618 polypeptide, useful as a vaccine for prophylactic and therapeutic treatment of cancers, particularly ovarian or colon cancer, autoimmune diseases and related conditions.

XX Claim 13; Page 61; 76pp; English.

XX The present sequence is the coding sequence of human CASB618 protein. The gene for human CASB618 is thought to be located on chromosome 15. The protein encoded by the present sequence and epitopes of the CASB618 protein (see AAB26327 to AAB26399) are useful in diagnosing the occurrence

CC of tumour cells and in vaccines for prophylactic and therapeutic
CC treatment of cancers, particularly ovarian or colon cancer, autoimmune
CC diseases and related conditions

```

SQ      Sequence 1441 BP; 289 A; 466 C; 392 G; 294 T; 0 U; 0 Other;

Query Match      100.0%; Score 1441; DB 3; Length 1441;
Best Local Similarity 100.0%; Pred. No. 1.3e-311;
Matches 1441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	1	AAAGTAA	CGGCTACAGACAGTGA	AAATAGTTTCGTGCGCGGCTAGAAAAA	CTCTGTCG	60
Db						
QY	1	AAAGTAA	CGGCTACAGACAGTGA	AAATAGTTTCGTGCGCGGCTAGAAAAA	CTCTGTCG	60
Db						
QY	61	GTACCAA	CCCCAGAGCGTTGAGAG	CAGCCCCACCTCCACGCTTCTTAA	CGGAGAGGTGCA	120
Db						
QY	61	GTACCAA	CCCCAGAGCGTTGAGAG	CAGCCCCACCTCCACGCTTCTTAA	CGGAGAGGTGCA	120
Db						
QY	121	GGACTCAGAC	TTTCAACGAGCCCACTCGGT	CCAGCCTTGTACGCA	AAAGAGACGCCAAGSAC	180
Db						
QY	121	GGACTCAGAC	TTTCAACGAGCCCACTCGGT	CCAGCCTTGTACGCA	AAAGAGACGCCAAGSAC	180
Db						
QY	181	GCGCTCT	CCCCGCTCCAGGCAG	CCCGCAGCTTGTCTGGCTTGCCTG	CCCGCCTGCGTG	240
Db						
QY	181	GCGCTCT	CCCCGCTCCAGGCAG	CCCGCAGCTTGTCTGGCTTGCCTG	CCCGCCTGCGTG	240
Db						
QY	241	ACTCGGCCG	GTGTCAGCATGAC	CCCTGTGGAA	CGGCGTACTGCC	300
Db						
QY	241	ACTCGGCCG	GTGTCAGCATGAC	CCCTGTGGAA	CGGCGTACTGCC	300
Db						
QY	301	CGGCATG	CCGACAGGCTTCAGCG	TTCCACTGCTCAT	CGTTATTCTAGTG	360
Db						
QY	301	CGGCATG	CCGACAGGCTTCAGCG	TTCCACTGCTCAT	CGTTATTCTAGTG	360
Db						
QY	361	GCAGCAAG	CTTCTGCTCATCTTGC	CGGGATCCGTTGGCCACTCG	CGCTGGTTTGGTTG	420
Db						
QY	361	GCAGCAAG	CTTCTGCTCATCTTGC	CGGGATCCGTTGGCCACTCG	CGCTGGTTTGGTTG	420
Db						
QY	421	GTGAGAG	TTCTTCTCAGTCTG	TTTCATAGGCGCAGAA	TTGTGGCTGTGCACTTCAG	480
Db						
QY	421	GTGAGAG	TTCTTCTCAGTCTG	TTTCATAGGCGCAGAA	TTGTGGCTGTGCACTTCAG	480
Db						
QY	481	GAATGTT	TCGTGGTACAGTGA	ACCAACACATCTCA	AAAGCCTTCAGCGCAGCGCGC	540
Db						
QY	481	GAATGTT	TCGTGGTACAGTGA	ACCAACACATCTCA	AAAGCCTTCAGCGCAGCGCGC	540
Db						
QY	541	GTACAGC	CCCGTCTGCTCGTGG	CGCTTGGAGGCGCATTA	ATATACACTCACAGGG	600
Db						
QY	541	GTACAGC	CCCGTCTGCTCGTGG	CGCTTGGAGGCGCATTA	ATATACACTCACAGGG	600
Db						
QY	601	ACCCAGT	GCATCAGCTGAA	CGAGACCATTGACT	ACAACGAGCAGTTCACTGGCGTCTG	660
Db						
QY	601	ACCCAGT	GCATCAGCTGAA	CGAGACCATTGACT	ACAACGAGCAGTTCACTGGCGTCTG	660
Db						
QY	661	AAAGAGA	ATTACGCGCGGAGT	ACGGAAACGCACTG	GAGAAAGGGCTGCCCGACCCAGT	720
Db						
QY	661	AAAGAGA	ATTACGCGCGGAGT	ACGGAAACGCACTG	GAGAAAGGGCTGCCCGACCCAGT	720
Db						
QY	721	CTCTAC	CTGGCGGAAAGTTCA	CCAGAGTAGCCCTTGG	CGCCTGTACCA	780
Db						
QY	721	CTCTAC	CTGGCGGAAAGTTCA	CCAGAGTAGCCCTTGG	CGCCTGTACCA	780
Db						
QY	781	CTGGCGG	GACACTACGCTCG	CGCCACGCTATGGG	TGGTGCTTCTGGCTCCCTCTCC	840
Db						
QY	781	CTGGCGG	GACACTACGCTCG	CGCCACGCTATGGG	TGGTGCTTCTGGCTCCCTCTCC	840
Db						
QY	841	AACGTG	CTGCTCCACGCG	CGCCCGCTCTAC	GAGGCGCTG	900
Db						
QY	841	AACGTG	CTGCTCCACGCG	CGCCCGCTCTAC	GAGGCGCTG	900
Db						
QY	901	GCGTTG	CGGCTCTTGGGGT	CTTGGCTTGGCCCT	CTATAGCGTGC	960
Db						
QY	901	GCGTTG	CGGCTCTTGGGGT	CTTGGCTTGGCCCT	CTATAGCGTGC	960
Db						

QY	961	CTCCGCGCTAGGCTCCTCGCGCTACCACTCAGTACGGCGCGCTTCTGGGTACAGCTG	1020
DB	961		1020
QY	1021	GCAACGGCGTCTGTGCTCTTCTCGAGGGCGGTGGTGAATCTCAGTATGTTCCG	1080
DB	1021		1080
QY	1081	CCCAGCGCTCTTCGCA CCGTTCTGGACCAAGCGCCAAAGACTGCAGCCAGGAGAGAGG	1140
DB	1081		1140
QY	1141	GGCTCACCTCTTATCCTCGGCGACCCACTGCAACAGCAGCGCCGCTCTCCAGACTTAAAA	1200
DB	1141		1200
QY	1201	TGTATCACCACTAACCTGTGAGGGGGACCCAATCTGGAATCTTCCCGCGCTTGGGACAT	1260
DB	1201		1260
QY	1261	CGCAGCGCGGAAGCAGTGC CCGCCAGGCTTGGGCCAGGAGAGCTCCAGGAAGGGCACTG	1320
DB	1261		1320
QY	1321	AGCGCTGTGGCGGAGGCTCGGACATCCG CAGGCAC CAGGGGAAAGTCTCTGGGGCGA	1380
DB	1321		1380
QY	1381	TCTGTAATAAACCCTTTTCTTTTGTGTTTTTAAAAA AAAAAAAAAAAAAAAAAA	1440
DB	1381		1440
QY	1441	A 1441	
DB	1441	A 1441	

RESULT 2

ADD19228

ID ADD19228 standard; cDNA; 1491 BP.

AC ADD19228:

DT 15-JAN-2004 (first entry)

Human cDNA from secreted protein gene 45:

XX human CD8a from secreted protein gene 15.
KW human secreted protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW nootropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; cardiovascular disorder; proliferative disorder; human; ss gene.

XX Homo sapiens.

XX PN WO2003052377-A2.

PD 26-JUN-2003.

12 20 JUN 2002;
XX
PF 06-NOV-2002; 2002MO-US035606.

PR 07-NOV-2001; 2001US-0331046P.

XX PA (HUMA-) HUMAN GENOME SCI INC.
07 NOV 2001 180200Z

PI Rosen CA, Ruben SM;

LI
XX
DB WPI: 2003-533050/50.

DR P-PSDB; ADD19303.

XX New isolated nucleic acids encoding signal transduction pathway component

PT polypeptides, useful for diagnosing, treating, and/or preventing

PT disorders, such as cancer, infections, cardiovascular and inflammatory

PT diseases.

XX Claim 1; SEQ ID NO 55; 554pp; English.

PS The invention relates to an isolated nucleic acid molecule (cDNA)

XX encoding a human secreted protein, representing one of 85 novel genes.

CC Also included are recombinant vectors, host cells (expressing the

CC protein), the secreted proteins (including their fragments, epitopes and

CC homologues), an isolated antibody that binds specifically to the protein,

CC diagnosing a pathological condition or susceptibility to a pathological

CC condition (comprising determining the presence or absence of a mutation

CC in the nucleic acid and diagnosing a condition based on the presence or

CC absence of the mutation), diagnosing a pathological condition or

CC susceptibility to a pathological condition (comprising determining the

CC presence or amount of expression of the protein in a biological sample

CC and diagnosing a condition based on the presence or amount of expression

CC of the protein), preventing, treating or ameliorating a medical condition

CC by administering the nucleic acid or protein to a mammalian subject,

CC identifying a binding partner to the protein, the gene corresponding to

CC the cDNA sequence, and identifying an activity in a biological assay

CC (comprising expressing the nucleic acid in a cell, isolating the

CC supernatant, detecting an activity in a biological assay and identifying

CC the protein in the supernatant having the activity). The nucleic acids

CC and proteins display the following activities: Cytostatic, antibacterial,

CC Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen,

CC Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-

CC Gen, Respiratory-Gen, Immunosuppressive, Cerebroprotective, Vasotropic,

CC Nootropic, Antiallergic. The methods and compositions of the present

CC invention are useful for diagnosing, treating, preventing and/or

CC prognosticating disorders related to the novel polypeptides, such as

CC cancer, bacterial or viral infections, and neural, immune system, blood,

CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,

CC renal, inflammatory or proliferative disorders (many examples of these

CC diseases and disorders are given in the specification). The present

CC sequence encodes a novel secreted protein of the invention.

XX Sequence 1491 BP; 293 A; 485 C; 410 G; 302 T; 0 U; 1 Other;

Query Match 99.7%; Score 1436; DB 9; Length 1491;

Best Local Similarity 99.9%; Pred. No. 1.7e-310;

Matches 1436; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGGTAGAAAACCTCTGTGC 60

DB 44 AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGGTAGAAAACCTCTGTGC 103

QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAAACGAGAGGTGCA 120

DB 104 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAAACGAGAGGTGCA 163

QY 121 GGAATCAGACTTACAGCCAGCCAGCTTCCAGCTTGTACGCAAGAGACGCCAAGGAC 180

DB 164 GGAATCAGACTTACAGCCAGCCAGCTTCCAGCTTGTACGCAAGAGACCGTCAAGGAC 223

QY 181 GCGCTCTCCCGCTCCAGGAGCCAGCTTGTGCTTGCCTTCCCTCCCGCTGCGTGCAGC 240

DB 224 GCGCTCTCCCGCTCCAGGAGCCAGCTTGTGCTTGCCTTCCCTCCCGCTGCGTGCAGC 283

QY 241 ACTCGGCGCGGCTGAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 300

DB 284 ACTCGGCGCGGCTGAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 343

QY 301 CGGCATGCCGAGGCTTCAGCGTTCCAGCTGCTCATGTTATCTAGTGTGTTTGGCTCTA 360

DB 344 CGGCATGCCGAGGCTTCAGCGTTCCAGCTGCTCATGTTATCTAGTGTGTTTGGCTCTA 403

QY 361 GCAGCAAGCTTCTGCTCATCTTCGCGGGGATCCGTGGCCACTCGCGCTGTTTGGTTG 420

DB 404 GCAGCAAGCTTCTGCTCATCTTTCGCGGGATCCGTGGCCACTCGCGCTGTTTGGTTG 463

QY 421 GTGAGAGTTCTTCTCAGTCTGTTCATAGGCGCAGAAATTTGTGGCTGTGCACTTTCAGTGCA 480

DB 464 GTGAGAGTTCTTCTCAGTCTGTTCATAGGCGCAGAAATTTGTGGCTGTGCACTTTCAGTGCA 523

QY 481 GAATGTTCTGTTGTTACAGTGAACACCAACACATCTTACAAAGCCCTTACGGCGCAGCGGC 540

DB 524 GAATGTTCTGTTGTTACAGTGAACACCAACACATCTTACAAAGCCCTTACGGCGCAGCGGC 583

QY 541 GTTACAGCCCGTGTCTGCTGTCTGTTGGGCTGTGAGGCGCATTAATTACACTCAGGCG 600

DB 584 GTTACAGCCCGTGTCTGCTGTCTGTTGGGCTGTGAGGCGCATTAATTACACTCAGGCG 643

QY 601 ACCCGAGTGCATCAGTGAACAGACCATTTGACTACACAGAGCAGTTTACCTTGGCTGCTG 660

DB 644 ACCCGAGTGCATCAGTGAACAGACCATTTGACTACACAGAGCAGTTTACCTTGGCTGCTG 703

QY 661 AAAGAGATTACCGCGGAGTACCGCAACGCACTGGAGAGGGGCTGCGGACCCAGTG 720

DB 704 AAAGAGATTACCGCGGAGTACCGCAACGCACTGGAGAGGGGCTGCGGACCCAGTG 763

QY 721 CTCTACCTGGGAGAGTTTACACCCAGTAGCCCTTGGGCTGTACACACAGTACAC 780

DB 764 CTCTACCTGGGAGAGTTTACACCCAGTAGCCCTTGGGCTGTACACACAGTACAC 823

QY 781 CTGGCGGACACTACCGCTCGGCGCACGCTATGTTGGTGGCTTCTGCTTCTGGCTCTCTCC 840

DB 824 CTGGCGGACACTACCGCTCGGCGCACGCTATGTTGGTGGCTTCTGCTTCTGGCTCTCTCC 883

QY 841 AACGTGCTCTCTCCACGCGCGCGCTCTACGGAGGCTGCGACTGTGACACACCGGA 900

DB 884 AACGTGCTCTCTCCACGCGCGCGCTCTACGGAGGCTGCGACTGTGACACACCGGA 943

QY 901 GCCTTCGCGCTCTTTCGCGCTCTTCGCTTGGCTTCCATCTCTAGCGTGGCGCTCTGCCG 960

DB 944 GCCTTCGCGCTCTTTCGCGCTCTTCGCTTGGCTTCCATCTCTAGCGTGGCGCTCTGCCG 1003

QY 961 CTCCGCTTAGGCTCTCTCCGCGCTCACCACTCAGTACGCGCGCGCTTCTGGGTTCAGCTG 1020

DB 1004 CTCCGCTTAGGCTCTCTCCGCGCTCACCACTCAGTACGCGCGCGCTTCTGGGTTCAGCTG 1063

QY 1021 GCAACCGGCTCTCTGCTCTTCTTCGAGGGGCGCTGTTGAGTCTCCAGTATGTTCCG 1080

DB 1064 GCAACCGGCTCTCTGCTCTTCTTCGAGGGGCGCTGTTGAGTCTCCAGTATGTTCCG 1123

QY 1081 CCCAGCGCTCTTCGACCCCTTCTGGACCAAGCGCCAGGACTGCGAGCAGGAGAGGG 1140

DB 1124 CCCAGCGCTCTTCGACCCCTTCTGGACCAAGCGCCAGGACTGCGAGCAGGAGAGGG 1183

QY 1141 GGCTCACCTCTTATTCCTCGGCGACCCACTGCAAGCAGGCGCTCTCCAGACTTAA 1200

DB 1184 GGCTCACCTCTTATTCCTCGGCGACCCACTGCAAGCAGGCGCTCTCCAGACTTAA 1243

QY 1201 TGTATCACCACTAACCTGTGAGGGGAGCCCAATCTGGAATCTTCCCGCTTGGGACAT 1260

DB 1244 TGTATCACCACTAACCTGTGAGGGGAGCCCAATCTGGAATCTTCCCGCTTGGGACAT 1303

QY 1261 CGCAGGCGGGAAGCAGTTCGCGCGCTGGCCAGGAGCTCCAGGAGGCGACTG 1320

DB 1304 CGCAGGCGGGAAGCAGTTCGCGCGCTGGCCAGGAGCTCCAGGAGGCGACTG 1363

QY 1321 AGCGTCTGCGCGGAGGCTCGGACATCCGAGGCAACGAGGAAAGTCTCTCTGGGCGGA 1380

DB 1364 AGCGTCTGCGCGGAGGCTCGGACATCCGAGGCAACGAGGAAAGTCTCTCTGGGCGGA 1423

QY 1381 TCTGTAAATAAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1438

DB 1424 TCTGTAAATAAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1481

RESULT 3

ABL90535

ID XX ABL90535 standard; cDNA; 1474 BP.
AC ABL90535;
XX
DT 24-MAY-2002 (first entry)
DE Human polynucleotide SEQ ID NO 1097.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR P-PSDB; ABB90126.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 4; SEQ ID NO 1097; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB9040-ABB9044) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1474 BP; 292 A; 480 C; 404 G; 296 T; 0 U; 2 Other;

Query Match 99.6%; Score 1435.6; DB 6; Length 1474;
Best Local Similarity 99.8%; Pred. No. 2.1e-310;
Matches 1435; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTGCGCGGCTAGAAAAAATCTGTGCG 60
DB
QY 27 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTGCGCGGCTAGAAAAAATCTGTGCG 86
DB
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGGAGAGGTGCA 120
DB
QY 87 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGGAGAGGTGCA 146
DB
QY 121 GGACTCAGACTTCACCGACCCACTCGGTCCAGCCTTGTACGCAAGAGACGCCAGGAC 180
DB 147 GGACTCAGACTTCACCGACCCACTCGGTCCAGCCTTGTACGCAAGAGACCGTCAAGGAC 206

QY 181 GCGCTCTCCGCTCCAGGCTCCAGGCTCCAGCTTGTGCTGGCTTGGCTGCGCGCTGCGTGCGAGC 240
DB 207 GCGCTCTCCGCTCCAGGCTCCAGGCTCCAGCTTGTGCTGGCTTGGCTGCGCGCTGCGTGCGAGC 266
QY 241 ACTCGGCGCGCTGCGAGCATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGAGCCC 300
DB 267 ACTCGGCGCGCTGCGAGCATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGAGCCC 326
QY 301 CGGCATGCGCGAGGCTTTCAGCGTCCACTGCTCATGCTTATTTAGTGTGCTGCTGCTA 360
DB 327 CGGCATGCGCGAGGCTTTCAGCGTCCACTGCTCATGCTTATTTAGTGTGCTGCTGCTA 386
QY 361 GCAGCAAGCTTCTGCTCATCTTCCGCGGATCCGCTGGCCACTGCGCGCTGCTGCTGCTG 420
DB 387 GCAGCAAGCTTCTGCTCATCTTCCGCGGATCCGCTGGCCACTGCGCGCTGCTGCTGCTG 446
QY 421 GTGAGAGTTCTTCTCAGTCTGTTTATAGGCGAGAAATTTGCTGCTGCTGCTGCTGCTGCA 480
DB 447 GTGAGAGTTCTTCTCAGTCTGTTTATAGGCGAGAAATTTGCTGCTGCTGCTGCTGCTGCA 506
QY 481 GAATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 507 GAATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566
QY 541 GTTACAGCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 567 GTTACAGCCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
QY 601 ACCCAGTGCATCAGCTGAACAGACCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 627 ACCCAGTGCATCAGCTGAACAGACCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 686
QY 661 AAAGAGAAATTAAGCGCGGAGTACCGGAAACCGACTGGAGAGGGGCTGCGGAGCCAGTG 720
DB 687 AAAGAGAAATTAAGCGCGGAGTACCGGAAACCGACTGGAGAGGGGCTGCGGAGCCAGTG 746
QY 721 CTCTACCTGCGGAGAAATTTACACCGAGTACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 747 CTCTACCTGCGGAGAAATTTACACCGAGTACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
QY 781 CTGCGGGGACACTACGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 807 CTGCGGGGACACTACGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
QY 841 AACGTGCTGCTTCCACCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 867 AACGTGCTGCTTCCACCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
QY 901 GCCTTCGCGCTTCTGCGGGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 927 GCCTTCGCGCTTCTGCGGGTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986
QY 961 CTCCGCTAGGCTTCTCCGCGCTCACCACCTCAGTACGCGCGCGCTTCTGCTGCTGCTGCTGCTG 1020
DB 987 CTCCGCTAGGCTTCTCCGCGCTCACCACCTCAGTACGCGCGCGCTTCTGCTGCTGCTGCTGCTG 1046
QY 1021 GCAACCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1047 GCAACCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
QY 1081 CCCAGCGCTTCTGCGCACCTTCTGGAACCAAGCGCAAGGACTGCGCGGAGAGAGGG 1140
DB 1107 CCCAGCGCTTCTGCGCACCTTCTGGAACCAAGCGCAAGGACTGCGCGGAGAGAGGG 1166
QY 1141 GCTCACTCTTATCTCGGAGCCCACTGCAAGCAGGCGCTCTCCAGACTTAAAA 1200
DB 1167 GCTCACTCTTATCTCGGAGCCCACTGCAAGCAGGCGCTCTCCAGACTTAAAA 1226
QY 1201 TGTATCACCACCTGAGGGGGACCCCAATCTGGAATCTGGAATCTCCCGCTTGGGACAT 1260
DB 1227 TGTATCACCACCTGAGGGGGACCCCAATCTGGAATCTGGAATCTCCCGCTTGGGACAT 1286

Db 1029 GCAACGGCGTCTGTGCTCTTCTCGAGGGCGGTGAGTCTCCAGTATGTCGG 1088
Qy 1081 CCCAGCGCTCTTGGACCTTCTGGACCAAGCGCCAGGACTGCGAGCAGGAGAGGG 1140
Db 1089 CCCAGCGCTCTTGGACCTTCTGGACCAAGCGCCAGGACTGCGAGCAGGAGAGGG 1148
Qy 1141 GGCTACCTCTTATCTCGGCGACCCACTGCACAAGCAGGCGCTCTCCAGACTTAAAA 1200
Db 1149 GGCTACCTCTTATCTCGGCGACCCACTGCACAAGCAGGCGCTCTCCAGACTTAAAA 1208
Qy 1201 TGTATCACCACTAACCTGTGAGGGGACCCAACTCTGGAATCTTCCCGCCTTGGGACAT 1260
Db 1209 TGTATCACCACTAACCTGTGAGGGGACCCAACTCTGGAATCTTCCCGCCTTGGGACAT 1268
Qy 1261 CGCAGGCGGGAAGCAGTGGCCCGCAGGCGCTGGCCAGGAGAGCTCCAGGAAGGGCACTG 1320
Db 1269 CGCAGGCGGGAAGCAGTGGCCCGCAGGCGCTGGCCAGGAGAGCTCCAGGAAGGGCACTG 1328
Qy 1321 AGCGCTGCTGGCGGAGGCTCGGACATCCGCGAGGACCCAGGAAAGTCTCTTGGGCGGA 1380
Db 1329 AGCGCTGCTGGCGGAGGCTCGGACATCCGCGAGGACCCAGGAAAGTCTCTTGGGCGGA 1388
Qy 1381 TCTGTAATAAACCTTTTCTTTTGTGTTTTT 1413
Db 1389 TCTGTAATAAACCTTTTCTTTGTTTTT 1421

RESULT 5
AAC95564
ID AAC95564 standard; cDNA; 1460 BP.
AC AAC95564;
XX
XX
DT 21-FEB-2001 (first entry)
DE Human secreted protein gene 44 SEQ ID NO:54.
XX
KW Human; secreted protein; cytostatic; immunosuppressive; neutrotropic;
KW neutrotropic; antiviral; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antiulcer; vulnery; anticonvulsant; antibacterial;
KW antifungal; antiparasitic; cardiant; cancer; immune disease; allergy;
KW cardiovascular disorder; wound healing; infection; neurological disease;
KW ss.
XX
OS Homo sapiens.
XX
XX WO200061596-A1.
XX
XX 19-OCT-2000.
XX
XX 06-APR-2000; 2000WO-US008983.
XX
XX 09-APR-1999; 99US-0128703P.
XX
XX 14-JAN-2000; 2000US-0176068P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Ruben SM, Komatsoulis G;
XX
XX WPI; 2000-611865/58.
XX
XX P-PSDB; AAB52055.
XX
XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
XX the prevention, treatment and diagnosis of cancer, immune disorders,
XX cardiovascular disorders and neurological diseases.
XX
XX Claim 1; Page 443; 505pp; English.
XX
XX Polynucleotide sequences AAC95521 - AAC95570 represent cDNA encoding 50
XX human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -
XX AAB52103 represent alternative polypeptides encoded by the genes, and

CC amino acid sequences with which they share homology. The genes and
CC proteins have activities dependent on the tissues and cells in which they
CC are expressed. Examples of their activities include cytostatic;
CC immunosuppressive; neutrotropic; neuroprotective; antiviral; antiallergic;
CC hepatotropic; antidiabetic; antiinflammatory; antiulcer; vulnery;
CC anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiant.
CC The secreted proteins, polypeptides, antagonists and agonists may be
CC useful in treating, preventing and/or diagnosing diseases and disorders
CC such as cancer, particularly breast and ovarian cancer, and other cancers
CC of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,
CC liver, lung, or urogenital. Immune disorders such as Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular diseases such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases such as viral,
CC bacterial, fungal and parasitic infections may also be treated using the
CC proteins and polynucleotides of the invention. Sequences AAC95512 -
CC AAC95520 and AAB52011 are used in the isolation and characterisation of
CC the proteins and polynucleotides of the invention
XX
SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;

Query Match 96.7%; Score 1392.8; DB 3; Length 1460;
Best Local Similarity 99.3%; Pred. No. 7.5e-301;
Matches 1430; Conservative 0; Mismatches 7; Indels 3; Gaps 3;

Qy 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGCGCTAGAAAACTCTGTG 60
Db 22 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGCGCTAGAAAACTCTGTG 81
Qy 61 GTACCAACCCCGAGCGTTGAGAGCAGCCACCTCCAGCTTCCTTAACGGAGAGGTGCA 120
Db 82 GTACCAACCCCGAGCGTTGAGAGCAGCCACCTCCAGCTTCCTTAACGGAGAGGTGCA 141
Qy 121 GGACTCAGACTTCACAGCCACTCGGTCCAGACCTTGTACCAAGAGAGCGCAAGGAC 180
Db 142 GGACTCAGACTTCACAGCCACTCGGTCCAGACCTTGTACCAAGAGAGCGCAAGGAC 201
Qy 181 GCGCTCTCCCGGTCCAGGAGCCCGAGCTTGGCTTGGCTGCCCGCTGGTGCAGC 240
Db 202 GCGCTCTCCCGGTCCAGGAGCCCGAGCTTGGCTTGGCTGCCCGCTGGTGCAGC 261
Qy 241 ACTCGGCGCGGTGCGAGCATCACCTGTGGAACGGCGTACTGCCCTTTTACCCCGAGCCC 300
Db 262 ACTCGGCGCGGTGCGAGCATCACCTGTGGAACGGCGTACTGCCCTTTTACCCCGAGCCC 321
Qy 301 CGGCATGCCGAGGCTTCAGGCTTCCACTGCTCATCGTTATCTAGTGTGTTGGCTCTA 360
Db 322 CGGCATGCCGAGGCTTCAGGCTTCCACTGCTCATCGTTATCTAGTGTGTTGGCTCTA 381
Qy 361 GCAGCAAGCTTCTGCTCATCTTCCCGGGATCCGTGGCGACTCGCGCTGGTTGGTTG 420
Db 382 GCAGCAAGCTTCTGCTCATCTTCCCGGGATCCGTGGCGACTCGCGCTGGTTGGTTG 441
Qy 421 GTGAGAGTCTTCTCAGTCTGTTTCATAGCGCGAGAAATTGTGGCTGTGCACCTCAGTGCA 480
Db 442 GTGAGAGTCTTCTCAGTCTGTTTCATAGCGCGAGAAATTGTGGCTGTGCACCTCAGTGCA 501
Qy 481 GAATGGTTCTGTTGGGTACAGTGAACACCAACACATCCTACAAAGCCTTCAGCGCAGCGCG 540
Db 502 GAATGGTTCTGTTGGGTACAGTGAACACCAACACATCCTACAAAGCCTTCAGCGCAGCGCG 561
Qy 541 GTTACAGCCCGTGTGGTCTGCTCGTGGGCTGGAGGCGCATTAATATACCTCAGCAGGG 600
Db 562 GTTACAGCCCGTGTGGTCTGCTCGTGGGCTGGAGGCGCATTAATATACCTCAGCAGGG 621
Qy 601 ACCCCAGTGCATCAGCTGAAGAGACCAATTGACTACACGAGCAGTTTACCTGGCGGTCTG 660
Db 622 ACCCCAGTGCATCAGCTGAAGAGACCAATTGACTACACGAGCAGTTTACCTGGCGGTCTG 681
Qy 661 AAAGAGAAATTACCGCGGAGTACGCGAACGCACTGGAGAGGGGCTGCGGAGCCAGTG 720

Db 682 AAGAGAAATTACGGCGGAGTACCGGAACGCACTGGAGAGGGGCTGCCGACCCAGTG 741
Qy 721 CTCTACCTGGCGGAGAAAGTTACACCGAGTAGCCTTGCGGCTGTACCACAGTACCAC 780
Db 742 CTCTACCTGGCGGAGAAAGTTACACCGAGTAGCCTTGCGGCTGTACCACAGTACCAC 801
Qy 781 CTGGCGGAGACATFACGCTCTCGGCGACGCTATGGGTGGCGTTCTGCTCTCTCTCTCC 840
Db 802 CTGGCGGAGACATFACGCTCTCGGCGACGCTATGGGTGGCGTTCTGCTCTCTCTCTCC 861
Qy 841 AACGTGCTCTCTCCACCGCGGCGGCGCTCTACGGAGGCTTGGCACTGTGACACCGGA 900
Db 862 AACGTGCTCTCTCCACCGCGGCGGCGCTCTACGGAGGCTTGGCACTGTGACACCGGA 921
Qy 901 GCCTTCGCGCTCTCTGGGGTCTTTCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 960
Db 922 GCCTTCGCGCTCTCTGGGGTCTTTCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 981
Qy 961 CTGGCGCTAGGCTCTCTCGGCGCTTACCACTACGAGGCGCGGCTTCTGGGTACGCTG 1020
Db 982 CTGGCGCTAGGCTCTCTCGGCGCTTACCACTACGAGGCGCGGCTTCTGGGTACGCTG 1041
Qy 1021 GCAACGGCGCTCTCTGCTCTCTCTCGGAGGGCGGCTGTGAGTCTCCAGTATGTTGG 1080
Db 1042 GCAACGGCGCTCTCTGCTCTCTCTCGGAGGGCGGCTGTGAGTCTCCAGTATGTTGG 1101
Qy 1081 CCAGGCGCTCTCTGCACTCTCTGACCAAGGCGGCAAGGCTGCG-AGCCAGAGAGAGG 1139
Db 1102 CCAGGCGCTCTCTGCACTCTCTGACCAAGGCGGCAAGGCTGCGAGCCAGAGAGAGG 1161
Qy 1140 GGGCTCACCTCTTATCTCTCGGCGGAGCCCACTGCACAGGAGGCGGCTCTCCAGAC-TTAA 1198
Db 1162 GGGCTCACCTCTTATCTCT-GGGGAGCCCACTGCACAGGAGGCGGCTTCCAGACTTTAA 1220
Qy 1199 AATGTATCACCACTAATCTGTGAGGGGAGCCCAATCTGGACTCTCTCCCGCTTGGGAC 1258
Db 1221 AATGTATCACCACTAATCTGTGAGGGGAGCCCAATCTGGACTCTCTCCCGCTTGGGAC 1280
Qy 1259 ATCCAGGCGGAGGAGCAGTGTCCCGCGGAGGCTGCGGCGGAGGAGCTCCAGGAGGGCAC 1318
Db 1281 ATCCAGGCGGAGGAGCAGTGTCCCGCGGAGGCTGCGGCGGAGGAGCTCCAGGAGGGCAC 1340
Qy 1319 TGAGCGCTGTGGCGGAGGCGCTCGGACATCCGAGGCGCACGAGGAAAGTCTCTGGGGC 1378
Db 1341 TGAGCGCTGTGGCGGAGGCGCTCGGACATCCGAGGCGCACGAGGAAAGTCTCTGGGGC 1400
Qy 1379 GATCTGTAATAAACCTTTTCTTTTCTTTTAAATAAAAAAAAAAAAAAAAAAAAA 1438
Db 1401 GATCTGTAATAAACCTTTTCTTTTCTTTTAAATAAAAAAAAAAAAAAAAAAAAA 1460

RESULT 6
ABZ67235
ID ABZ67235 standard; cDNA; 1460 BP.
XX
AC ABZ67235;
XX
DT 26-MAR-2003 (first entry)
XX
DE Human secreted protein encoding cDNA SEQ ID NO 355.
XX
KW Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnery; antibacterial; antiparkinsonian; antischlicking; antianaemic;
KW antianthratic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiac; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; chromosome 9p21; ds.
OS Homo sapiens.
XX
PN WO200277186-A2.

XX 03-OCT-2002.
PD
XX 26-MAR-2002; 2002WO-US009188.
PF
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
PI WPI; 2003-040583/03.
XX P-PSDB; ABP99814.
DR New human secreted proteins encoded by genes contained in cDNA clones
DR (e.g. HGCA19), useful for preventing, treating or diagnosing e.g. AIDS,
XX multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
XX West Nile fever.
PS Claim 7; Page 1353; 2423pp; English.
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
XX
SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;
Query Match 96.7%; Score 1392.8; DB 7; Length 1460;
Best Local Similarity 99.3%; Pred. No. 7.5e-301;
Matches 1430; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
Qy 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCCCGGCTAGAAAACTCTGTGC 60
Db 22 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCCCGGCTAGAAAACTCTGTGC 81
Qy 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGAGAGAGGTGCA 120
Db 82 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGAGAGAGGTGCA 141
Qy 121 GGACTCAGACTTCACCGCCCACTCGGTCCCGAGCTTGTACGCAAGAGAGAGGAGGAC 180
Db 142 GGACTCAGACTTCACCGCCCACTCGGTCCCGAGCTTGTACGCAAGAGAGAGGAGGAC 201
Qy 181 GCGCTCTCCCGGCTCCAGCAGCCCGGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTG 240
Db 202 GCGCTCTCCCGGCTCCAGCAGCCCGGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTG 261
Qy 241 ACTCGGCGCGGCTGCGAGCATGACCCCTGTGGAACGGCGTACTGCTTTCACCCCGGAGCC 300
Db 262 ACTCGGCGCGGCTGCGAGCATGACCCCTGTGGAACGGCGTACTGCTTTCACCCCGGAGCC 321
Qy 301 CGGCATGCGCAGGCTTTCAGCGCTTCCAGCTCATCGTATCTAGTGTGTTTGGCTCTA 360
Db 322 CGGCATGCGCAGGCTTTCAGCGCTTCCAGCTCATCGTATCTAGTGTGTTTGGCTCTA 381
Qy 361 GCAGCAAGCTTCTGCTCATCTTTCGCGGGGATCCGTCGCGCATCTCGCGCTGTTTGGTTG 420
Db 382 GCAGCAAGCTTCTGCTCATCTTTCGCGGGGATCCGTCGCGCATCTCGCGCTGTTTGGTTG 441

[illegible]

QY	1199	AATGTATCACCACTAACTGTGAGGGGACCCAAATCTGGACTCTTCCCGGCTTGGGAC	1258
Db	1221	AATGTATCACCACTAACTGTGAGGGGACCCAAATCTGGACTCTTCCCGGCTTGGGAC	1280
QY	1259	ATCGAGGCCGGGAAGCAGTGTCCCGCCAGGCTGGGCCAGGAGAGCTCCAGGAAGGGCAC	1318
Db	1281	ATCGAGGCCGGGAAGCAGTGTCCCGCCAGGCTGGGCCAGGAGAGCTCCAGGAAGGGCAC	1340
QY	1319	TGAGCGCTGTGTGGCGGAGGCTCGGACATCCCGAGGCCACGAGGAAAGTCTCTCGGGC	1378
Db	1341	TGAGCGCTGTGTGGCGGAGGCTCGGACATCCCGAGGCCACGAGGAAAGTCTCTCGGGC	1400
QY	1379	GATCTGTAATAAACCTTTTCTTTCTTTTAAAAAATAAAAAAAAAAAAAAAAAA	1438
Db	1401	GATCTGTAATAAACCTTTTCTTTCTTTTAAAAAATAAAAAAAAAAAAAAAAAA	1460

RESIST

RESULT 8
ADC20289

ADC20289
ID ADC20289 standard: DNA: 1460 bp.

AC ADC20289;

DT 18-DEC-2003 (first entry)

XX Human secreted protein coding sequence #228.

gene therapy; human; secreted protein; haemopoietic disorder;
haematological disorder; anaemia; haemophilia; inflammatory disorder;
inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
leukaemia; wound healing; epithelial cell proliferation disorder;
immune disorder; autoimmune disorder; asthmatic disorder;
cardiovascular disorder; atherosclerosis; myocarditis;
infectious disease; HIV; AIDS; endocrine disorder; diabetes;
gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.

XX Homo sapiens.

XX PN WC200292787-A2

21-NOV-2002

26-MAR-2002: 2002WO-US009257.

XX
PR 27-MAR-2001: 2001US-0278650P.

PR 12-SEP-2001; 2001US-0095008Z;
PR 12-SEP-2001; 2001US-0095008Z;
PR 12-SEP-2001; 2001US-0095008Z;

XX
BA (HITM2-) HUMAN GENOME SCT INC

XX	Poseen CB	Ruben SM.
BT		

XX
WDT: 3003-100000-176

XX DE

PT preparing a diagnostic or pharmaceutical composition for diagnosing
PT preventing or treating hematopoietic or hematologic disorders, e.g.
PT anemia or hemophilia.

PS Claim 1; SEQ ID NO 238; 1512pp; English.

The invention comprises the amino acid and coding sequences of human secreted proteins. The DNA and protein sequences of the invention are useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating: haematopoietic or haematological disorders (e.g. anaemia and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia); wound healing and disorders of epithelial cell proliferation; immune disorders (e.g. autoimmune disorders and asthmatic disorders); cardiovascular disorders (e.g. atherosclerosis and myocarditis); infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes); and gastrointestinal disorders (e.g. duodenal ulcers and gastroenteritis). The present DNA sequence encodes a human secreted

CC cancers, particularly ovarian or colon cancer, autoimmune diseases and
CC related conditions
XX
SQ Sequence 498 BP; 97 A; 166 C; 137 G; 98 T; 0 U; 0 Other;

Query Match 32.3%; Score 465.4; DB 3; Length 498;
Best Local Similarity 99.4%; Pred. No. 4.4e-94;
Matches 488; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 939 CTCTAGCGTCCGCTCTGCCGCT-CGCGCTAGGCTCCTCCGCGTCCACCTCAGTACG 997
Dd 1 CTCTAGCGTCCGCTCTGCCGCTCCTCCGCGTCCACCTCAGTACG 60
QY 998 -GCGCGCCTTCTGGGTACGCTGCAACCGGCGTCTGTGCTCTCTCTCGAGGGGC 1056
Dd 61 AGCGCGCCTTCTGGGTACGCTGCAACCGGCGTCTGTGCTCTCTCTCGAGGGGC 120
QY 1057 GTGTGAGTCTCCAGTATGTTCCGCCAGCGCTCTTCCGACCTTCTGGACCAAGCGCC 1116
Dd 121 GTGTGAGTCTCCAGTATGTTCCGCCAGCGCTCTTCCGACCTTCTGGACCAAGCGCC 180
QY 1117 AAGGACTGCAGCCAGGAGAGAGGGGCTCACCTTATCTCCTCGGACCCACCTGCACAAG 1176
Dd 181 AAGGACTGCAGCCAGGAGAGAGGGGCTCACCTTATCTCCTCGGACCCACCTGCACAAG 240
QY 1177 CAGGCGCTCTCCAGACTTAAATGTATCACCACTAACCTGTGAGGGGACCCCAATCTG 1236
Dd 241 CAGGCGCTCTCCAGACTTAAATGTATCACCACTAACCTGTGAGGGGACCCCAATCTG 300
QY 1237 GACTCTCTCCCGCTTGGGACATCGCAGGCGCGGAAGCAGTGGCCGCCAGGCTGGGC 1296
Dd 301 GACTCTCTCCCGCTTGGGACATCGCAGGCGCGGAAGCAGTGGCCGCCAGGCTGGGC 360
QY 1297 AGGAGAGCTCCAGGAAGGGCCTGAGCGCTGTGCGCGCGAGGCTCGGACATCCGAGGC 1356
Dd 361 AGGAGAGCTCCAGGAAGGGCCTGAGCGCTGTGCGCGCGAGGCTCGGACATCCGAGGC 420
QY 1357 ACCAGGGAAGTCTCCTGGGCGAFTCTGTAAATAAACCTTTTTTCTTTTAA 1416
Dd 421 ACCAGGGAAGTCTCCTGGGCGAFTCTGTAAATAAACCTTTTTTCTTTTAA 480
QY 1417 AAAAAAATAA 1427
Dd 481 AAAAAAATAA 491

RESULT 10
ABZ68115
ID ABZ68115 standard; DNA; 5033 BP.
XX ABZ68115;
AC
XX
DT 26-MAR-2003 (first entry)
XX
DE Human secreted protein encoding genomic DNA SEQ ID NO 1638.
XX
KW Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnary; antibacterial; antiparkinsonian; antiskinning; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antiulcer; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277186-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-US009188.
XX

PR 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-040583/03.
XX
PT New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGCA19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX
PS Disclosure; Page 2263-2264; 2423pp; English.
XX
CC The invention relates to novel human genes (ABZ6891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
XX
SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match 28.1%; Score 405.4; DB 7; Length 5033;
Best Local Similarity 98.6%; Pred. No. 2e-80;
Matches 409; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAAGTAACGGCTACAGACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Dd 13 AAAGTAACGGCTACAGACAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 72
QY 61 GTACCAACCCAGAGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Dd 73 GTACCAACCCAGAGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132
QY 121 GGACTCAGACTTACCAG 180
Dd 133 GGACTCAGACTTACCAG 192
QY 181 GCGCTCTCCCGGTCCAG 240
Dd 193 GCGCTCTCCCGGTCCAG 252
QY 241 ACTCGGCGCGGTGAG 300
Dd 253 ACTCGGCGCGGTGAG 312
QY 301 CGGCATGCGCAGAGGCTTCCAGGTTCCAGTGTCTATCTAGTGTGTGTGTGTGTGT 360
Dd 313 CGGCATGCGCAGAGGCTTCCAGGTTCCAGTGTCTATCTAGTGTGTGTGTGTGTGT 372
QY 361 GCAGCAAGCTTCTCTGCTCATCTTCCCGGAGATCCGAGGAGAGAGAGAGAGAGAG 415
Dd 373 GCAGCAAGCTTCTCTGCTCATCTTCCCGGAGATCCGAGGAGAGAGAGAGAGAGAG 427

RESULT 11
ABZ74587
ID ABZ74587 standard; DNA; 5033 BP.
XX
AC ABZ74587;

XX	DT	12-MAY-2003	(first entry)	
XX	DE	Secreted protein gene 360 genomic fragment HUFCL31, SEQ ID NO:1734.		
XX	KW	Human; secreted protein; cancer; tumour; hyperproliferative disorder; autoimmune disorder; inflammation; angiogenic diseases; AIDS; acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing; drug screening; chromosome identification; chromosome mapping; cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV; antianaemic; vulnary; chromosome 9p21; gene; ds.		
OS	OS	Homo sapiens.		
XX	PN	WO200277013-A2.		
XX	PD	03-OCT-2002.		
XX	PF	26-MAR-2002; 2002WO-US009370.		
XX	PR	27-MAR-2001; 2001US-0278650P.		
PR	PR	12-SEP-2001; 2001US-00950082.		
PR	PR	12-SEP-2001; 2001US-00950083.		
XX	PA	(HUMA-) HUMAN GENOME SCI INC.		
XX	PI	Rosen CA, Ruben SM;		
XX	XX	WPI; 2003-040578/03.		
XX	PT	New human secreted proteins and nucleic acids, useful for detecting or treating cancer or other hyperproliferative disorders, autoimmune disorders, inflammatory disorders, HIV disease, hepatitis or anemia.		
XX	PS	Disclosure; Page 2315-2316; 2474pp; English.		
XX	CC	ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted protein genes, and ABP00947-ABP01363 represent the proteins they encode.		
CC	CC	ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The invention also encompasses antibodies specific for the secreted proteins, the use of the secreted proteins in drug screening and recombinant vectors and host cells comprising a nucleic acid of the invention. The secreted proteins are thought to be involved in biological activities associated with cellular signalling, cellular differentiation, cell migration, prothorone activation and neurotransmitter activity. The secreted proteins, nucleic acids encoding them, antibodies or antibody fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing or treating cancers or other hyperproliferative disorders. Additionally, the secreted proteins and their nucleic acids may also be used in the treatment of autoimmune disorders, inflammatory disorders, diseases involving angiogenesis, AIDS (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, and as molecular weight markers. The present sequence represents a human secreted protein genomic fragment referred to in the disclosure of the invention		
XX	SQ	Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;		
Query Match		28.1%; Score 405.4; DB 7; Length 5033;		
Best Local Similarity		98.6%; Pred. No. 2e-80;		
Matches	409;	Conservative	0;	Mismatches
		Indels	0;	Gaps
QY	1	AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTCG 60		
DB	13	AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTCG 72		
QY	61	GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGGAGAGGTGCA 120		
DB	73	GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGGAGAGGTGCA 132		

QY	121	GGACTCAGACTTCACAGCCCACTCGGTCCAGCCTTGTATCGCAAGAGAGACGCCAAGGAC	180
DB	133	GGACTCAGACTTCACAGCCCACTCGGTCCAGCCTTGTATCGCAAGAGAGACGCCAAGGAC	192
QY	181	GGCTCTCCCGGTCCAGGAGCCAGCTTGTGCTGGCTTGCCTGCCCGCTGCTGCAGC	240
DB	193	GGCTCTCCCGGTCCAGGAGCCAGCTTGTGCTGGCTTGCCTGCCCGCTGCTGCAGC	252
QY	241	ACTCGGCGGGGTGCAGCATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGAGCCC	300
DB	253	ACTCGGCGGGGTGCAGCATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGAGCCC	312
QY	301	CGGCATGCCGCGAGGCTTCAGCGTTCCACTGCTCATCTGTTATCTAGTGTGCTCTA	360
DB	313	CGGCATGCCGCGAGGCTTCAGCGTTCCACTGCTCATCTGTTATCTAGTGTGCTCTA	372
QY	361	GCAGCAAGCTTCTGCTCATCTTGC CGGGGATCCGTCGCGCTGCTGCTTTT 415	
DB	373	GCAGCAAGCTTCTGCTCATCTTGC CGGGGATCCGTCGCGCTGCTGCTTTT 427	

RESULT 12

ADC21005

ID ADC21005 standard; DNA; 5033 BP.

XX AC ADC21005;

XX DT 18-DEC-2003 (first entry)

XX DE Human secreted protein-related DNA sequence #423.

XX KW gene therapy; human; secreted protein; haemopoietic disorder; haematological disorder; anaemia; haemophilia; inflammatory disorder; inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer; leukaemia; wound healing; epithelial cell proliferation disorder; immune disorder; autoimmune disorder; asthmatic disorder; cardiovascular disorder; atherosclerosis; myocarditis; infectious disease; HIV; AIDS; endocrine disorder; diabetes; gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.

XX OS Homo sapiens.

XX PN WO200292787-A2.

XX PD 21-NOV-2002.

XX PF 26-MAR-2002; 2002WO-US009257.

XX PR 27-MAR-2001; 2001US-0278650P.

XX PR 12-SEP-2001; 2001US-00950082.

XX PR 12-SEP-2001; 2001US-00950083.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX PT WPI; 2003-129287/12.

XX PT New human secreted proteins and nucleic acid molecules, useful for preparing a diagnostic or pharmaceutical composition for diagnosing, preventing or treating hematopoietic or hematologic disorders, e.g. anemia or hemophilia.

XX PS Disclosure; SEQ ID NO 959; 1512pp; English.

XX CC The invention comprises the amino acid and coding sequences of human secreted proteins. The DNA and protein sequences of the invention are useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating: haematopoietic or haematological disorders (e.g. anaemia and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia); wound healing and disorders of epithelial cell proliferation; immune disorders (e.g. autoimmune disorders and asthmatic disorders);

CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC and gastrointestinal disorders (e.g. duodenal ulcers and
CC gastroenteritis). The present DNA sequence was used in the
CC exemplification of the invention.
XX
SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;
Query Match 28.1%; Score 405.4; DB 9; Length 5033;
Best Local Similarity 98.6%; Pred. No. 2e-80;
Matches 409; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 AAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGTAGAAAACTCTGTGC 60
Db |||||
QY 13 AAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGTAGAAAACTCTGTGC 72
Db |||||
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGGAGGTGCA 120
Db |||||
QY 73 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAACGGAGGTGCA 132
QY 121 GGAATCAGACTTACAGAGCCCACTCGGTCCAGCTTGTACGCAAGAGAGCGCAAGGAC 180
Db |||||
QY 133 GGAATCAGACTTACAGAGCCCACTCGGTCCAGCTTGTACGCAAGAGAGCGCAAGGAC 192
QY 181 GCGCTCTCCCGTCCAGGCAGCCCGAGCTTGTGCTGCTTGCCTCCGCTGCGTGCAGC 240
Db |||||
QY 193 GCGCTCTCCCGTCCAGGCAGCCCGAGCTTGTGCTGCTTGCCTCCGCTGCGTGCAGC 252
QY 241 ACTCGCCGGCGTGAGCATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 300
Db |||||
QY 253 ACTCGCCGGCGTGAGCATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 312
QY 301 CGGATGCCGCGAGGCTTCCAGCTTCCAGCTTGTGCTGCTTGCCTTATTCAGTGTGTTT 360
Db |||||
QY 313 CGGATGCCGCGAGGCTTCCAGCTTCCAGCTTGTGCTGCTTGCCTTATTCAGTGTGTTT 372
QY 361 GCAGCAAGCTTCTGCTCATCTTCCGCGGATCCGTGCGCCACTCGGCTGTTT 415
Db |||||
QY 373 GCAGCAAGCTTCTGCTCATCTTCCGCGGATCCGTGCGCCACTCGGTAAGGTGT 427

RESULT 13
ADD19187
ID ADD19187 standard; cDNA; 580 BP.
AC ADD19187;
XX
XX
DT 15-JAN-2004 (first entry)
XX
XX Human cDNA from secreted protein gene 4.
DE
XX human secreted protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW nootropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; human; ss; gene.
XX
OS Homo sapiens.
XX
XX
PN WO2003052377-A2.
XX
PD 26-JUN-2003.
XX
XX
PF 06-NOV-2002; 2002WO-US035606.
XX
XX
PR 07-NOV-2001; 2001US-0331046P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX

PI Rosen CA, Ruben SM;
XX WPI; 2003-533050/50.
DR P-PSDB; ADD19262.
XX
PT New isolated nucleic acids encoding signal transduction pathway component
PT polypeptides, useful for diagnosing, treating, and/or preventing
PT disorders, such as cancer, infections, cardiovascular and inflammatory
PT diseases.
XX
PS Claim 1; SEQ ID NO 14; 554pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human secreted protein, representing one of 85 novel genes.
CC Also included are recombinant vectors, host cells (expressing the
CC protein), the secreted proteins (including their fragments, epitopes and
CC homologues), an isolated antibody that binds specifically to the protein,
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition (comprising determining the presence or absence of a mutation
CC in the nucleic acid and diagnosing a condition based on the presence or
CC absence of the mutation), diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or amount of expression of the protein in a biological sample
CC and diagnosing a condition based on the presence or amount of expression
CC of the protein), preventing, treating or ameliorating a medical condition
CC by administering the nucleic acid or protein to a mammalian subject,
CC identifying a binding partner to the protein, the gene corresponding to
CC the cDNA sequence, and identifying an activity in a biological assay
CC (comprising expressing the nucleic acid in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity). The nucleic acids
CC and proteins display the following activities: Cytostatic, antibacterial,
CC Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen,
CC Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-
CC Gen, Respiratory-Gen, Immunosuppressive, Cerebroprotective, Vasotropic,
CC Nootropic, Antiallergic. The methods and compositions of the present
CC invention are useful for diagnosing, treating, preventing and/or
CC prognosticating disorders related to the novel polypeptides, such as
CC cancer, bacterial or viral infections, and neural, immune system, blood,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, inflammatory or proliferative disorders (many examples of these
CC diseases and disorders are given in the specification). The present
CC sequence encodes a novel secreted protein of the invention.
XX
SQ Sequence 580 BP; 115 A; 198 C; 156 G; 111 T; 0 U; 0 Other;
Query Match 28.1%; Score 405; DB 9; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.4e-80;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGTAGAAAACTCTGTGC 60
Db |||||
QY 13 AAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGTAGAAAACTCTGTGC 72
Db |||||
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAACGGAGAGGTGCA 120
Db |||||
QY 73 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAACGGAGAGGTGCA 132
QY 121 GGAATCAGACTTACAGAGCCCACTCGGTCCAGCTTGTACGCAAGAGAGCGCAAGGAC 180
Db |||||
QY 133 GGAATCAGACTTACAGAGCCCACTCGGTCCAGCTTGTACGCAAGAGAGCGCAAGGAC 192
QY 181 GCGCTCTCCCGTCCAGGCAGCCCGAGCTTGTGCTGCTTGCCTCCGCTGCGTGCAGC 240
Db |||||
QY 193 GCGCTCTCCCGTCCAGGCAGCCCGAGCTTGTGCTGCTTGCCTCCGCTGCGTGCAGC 252
QY 241 ACTCGCCGGCGTGAGCATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 300
Db |||||
QY 253 ACTCGCCGGCGTGAGCATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCGAGCCC 312
QY 301 CGGATGCCGCGAGGCTTCCAGCTTCCAGCTTGTGCTGCTTGCCTTATTCAGTGTGTTT 360
Db |||||
QY 313 CGGATGCCGCGAGGCTTCCAGCTTCCAGCTTGTGCTGCTTGCCTTATTCAGTGTGTTT 372

QY 361 GCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCCACTCG 405
DB 373 GCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCCACTCG 417

RESULT 14
AAS66563/c
ID AAS66563 standard; cDNA; 406 BP.

AC AAS66563;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #2367.

XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR P-PSDB; ABG02376.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 1; SEQ ID NO 2367; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridization probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have application of mutations
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 406 BP; 83 A; 107 C; 138 G; 78 T; 0 U; 0 Other;

Query Match 27.9%; Score 401.8; DB 5; Length 406;
Best Local Similarity 99.5%; Pred. No. 6.4e-80;
Matches 403; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAAGTAACGGCTACAGACAGTGAAGATAGTTTCGTCGGCGGCTAGAAAAACTCTGTCG 60

DB 405 AAAGTAACGGCTACAGACAGTGAAGATAGTTTCGTCGGCGGCTAGAAAAACTCTGTCG 346
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGAGAGGTGCA 120
DB 345 GGACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGAGAGGTGCA 286
QY 121 GGACTCAGACTTTCACCAAGCCCACTCGGTCCAGCCTTGTACGAAAGAGACGCCAAGGAC 180
DB 285 GGACTCAGACTTTCACCAAGCCCACTCGGTCCAGCCTTGTACGAAAGAGACGTCGAGGAC 226
QY 181 GCGCTCTCCCGGCTCCAGGCAGCCCAAGCTTGTGCTGCTTGCCTGCTCCCGCTGCGTGCAGC 240
DB 225 GCGCTCTCCCGGCTCCAGGCAGCCCAAGCTTGTGCTGCTTGCCTGCTCCCGCTGCGTGCAGC 166
QY 241 ACTCGGCGCGGTGAGCATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCCAGCCC 300
DB 165 ACTCGGCGCGGTGAGCATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCCAGCCC 106
QY 301 CGGCATGCGCGCAGGCTTCAGCGTTCCACTGCTCATGCTTATTCTAGTGTGCTCTA 360
DB 105 CGGCATGCGCGCAGGCTTCAGCGTTCCACTGCTCATGCTTATTCTAGTGTGCTCTA 46
QY 361 GCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCCACTCG 405
DB 45 GCAGCAAGCTTCTGCTCATCTTGGCGGGGATCCGTGGCCACTCG 1

RESULT 15

ADA53627/c

ID ADA53627 standard; cDNA; 2684 BP.

XX ADA53627;

XX 20-NOV-2003 (first entry)

XX Human coding sequence, SEQ ID 1195.

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.

XX Homo sapiens.

XX EP1293569-A2.

XX 19-MAR-2003.

XX 21-MAR-2002; 2002EP-00006586.

XX 14-SEP-2001; 2001JP-00328381.

PR 24-JAN-2002; 2002US-0350435P.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-395539/38.

DR P-PSDB; ADA55266.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 1; SEQ ID NO 1195; 205pp; English.

XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,

CC inflammatory diseases, osteoporosis or neurological disease.

XX

SQ Sequence 2684 BP; 599 A; 785 C; 684 G; 616 T; 0 U; 0 Other;

Query Match 27.5%; Score 396; DB 7; Length 2684;

Best Local Similarity 97.6%; Pred. No. 2.1e-78;

Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy	1007	TCTGGGTACAGCTGGGCAACCGGCTCCTGTGCTCTCTCTCGAGGGGCGGTGAGTC	1066
Db	2363	TTTCGATCCCCACCGCCACAGGCTCTCTGCTCTCTCTCGAGGGGCGGTGAGTC	2304
Qy	1067	TCCAGTATGTTCCGCCCCAGCGCTCTTCGCACCCCTCTGGACCAAGCGCCAAAGACTGCA	1126
Db	2303	TCCAGTATGTTCCGCCCCAGCGCTCTTCGCACCCCTCTGGACCAAGCGCCAAAGACTGCA	2244
Qy	1127	GCCAGGAGAGAGGGGCTCACCTCTTATCCTCGCGGACCCACTGCACAGCAGGCCGCTC	1186
Db	2243	GCCAGGAGAGAGGGGCTCACCTCTTATCCTCGCGGACCCACTGCACAGCAGGCCGCTC	2184
Qy	1187	TCCAGACTTAAATGTATCACCACTAACCTGTGAGGGGACCCCAATCTGGACTCCTTCC	1246
Db	2183	TCCAGACTTAAATGTATCACCACTAACCTGTGAGGGGACCCCAATCTGGACTCCTTCC	2124
Qy	1247	CCGCTTGGGACATCGCAGGCCCGGAGCAGTGCCTCGGCGGAGGCTGGCCAGGAGCTC	1306
Db	2123	CCGCTTGGGACATCGCAGGCCCGGAGCAGTGCCTCGGCGGAGGCTGGCCAGGAGCTC	2064
Qy	1307	CAGGAAGGGCACTGAGCGCTGCTGGCGGAGGCTCGGACATCCGAGGCGACCCAGGAAA	1366
Db	2063	CAGGAAGGGCACTGAGCGCTGCTGGCGGAGGCTCGGACATCCGAGGCGACCCAGGAAA	2004
Qy	1367	GTCTCCTGGGGCGATCTGTAAATAAACCTTTTCTTTTCTTTTAAATAA	1418
Db	2003	GTCTCCTGGGGCGATCTGTAAATAAACCTTTTCTTTTCTTTTAAATAA	1952

Search completed: June 3, 2004, 10:20:49

Job time : 646 secs

Sequence 5066, Ap
Sequence 5127, Ap
Sequence 5094, Ap
Sequence 13572, A
Sequence 13773, A
Sequence 13656, A
Patent No. 5223425
Patent No. 5223425
Sequence 25, Appl
Sequence 548, Appl
Sequence 19, Appl
Sequence 446, Appl
Sequence 5564, Ap
Sequence 5601, Ap
Sequence 7, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli

28 43.4 3.0 597 4 US-09-252-991A-5066
29 43.4 3.0 1308 4 US-09-252-991A-5127
30 43.4 3.0 1782 4 US-09-252-991A-5094
31 43.2 3.0 561 4 US-09-252-991A-13572
32 43.2 3.0 4131 4 US-09-252-991A-13773
33 43.2 3.0 8211 4 US-09-252-991A-13656
34 43 3.0 835 6 5223425-7
35 43 3.0 1093 6 5223425-3
36 43 3.0 2089 4 US-09-291-922-25
37 43 3.0 2196 4 US-09-673-395A-548
38 42.8 3.0 1454 4 US-09-372-422A-19
39 42.6 3.0 367 3 US-09-328-111-446
40 42.6 3.0 744 4 US-09-252-991A-5564
41 42.6 3.0 1332 4 US-09-252-991A-5601
42 42.6 3.0 1681 4 US-09-434-288-7
43 42.4 2.9 68750 3 US-09-335-409-1
44 42.4 2.9 68750 4 US-09-568-102-1
45 42.4 2.9 68750 4 US-09-567-969-1

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14
Query Match 4.2%; Score 61; DB 1; Length 7218;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 3, 2004, 10:05:39 ; Search time 124 Seconds
(without alignments)
6449.066 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggtacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	61	4.2	7218	1	Sequence 14, Appl
2	50.2	3.5	1098	4	Sequence 1304, Ap
3	50.2	3.5	1257	4	Sequence 1255, Ap
4	50.2	3.5	1746	4	Sequence 1163, Ap
5	48.2	3.3	396	4	Sequence 1350, Ap
6	48	3.3	1436	4	Sequence 35, Appl
7	47.6	3.3	420	4	Sequence 11029, A
8	47.6	3.3	1683	4	Sequence 11226, A
9	47.6	3.3	1974	4	Sequence 11111, A
10	47.6	3.3	2283	4	Sequence 10956, A
11	46.4	3.2	1431	4	Sequence 49, Appl
12	45	3.1	1926	4	Sequence 4, Appli
13	45	3.1	1931	2	Sequence 2, Appli
14	44.4	3.1	1835	4	Sequence 16, Appl
15	44.4	3.1	152331	3	Sequence 1610, Ap
16	43.8	3.0	462	4	Sequence 1556, Ap
17	43.8	3.0	1362	4	Sequence 12931, A
18	43.8	3.0	1551	4	Sequence 1561, Ap
19	43.8	3.0	1587	4	Sequence 13081, A
20	43.8	3.0	1659	4	Sequence 1465, Ap
21	43.8	3.0	1917	4	Sequence 12858, A
22	43.8	3.0	2073	4	Sequence 2, Appli
23	43.8	3.0	5852	1	Sequence 9, Appli
24	43.6	3.0	3722	4	Sequence 5, Appli
25	43.6	3.0	3862	4	Sequence 7, Appli
26	43.6	3.0	3937	4	Sequence 3, Appli
27	43.6	3.0	3985	4	Sequence 3, Appli

Best Local Similarity 1.9%; Pred. No. 5.5e-05;
Matches 7; Conservative 221; Mismatches 131; Indels 0; Gaps 0;
QY 753 CCTTCGGCTTCTGCTTCTGCTCTCTCCACGTCGTCTCTCCACGCGCGCCGCTCTA 872
Db 1084 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1143
QY 813 GGTGGGCTTCTGCTTCTGCTCTCTCCACGTCGTCTCTCCACGCGCGCCGCTCTA 872
Db 1144 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1203
QY 873 OGGAGGCTGGCACTGCTGACCAACCGGAGCCTTCGGGCTCTTCGGGCTCTTCGGCTTGGC 932
Db 1204 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1263
QY 933 CTCCATCTCTAGCTGCTGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCTGCGGCTCA 992
Db 1264 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1323
QY 993 GTACGGCGCGGCTTCTGCTGCTCACTGCGGCTGCGGCTGCGGCTGCTGCTGCTGCTGCGG 1052
Db 1324 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1383
QY 1053 GGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1111
Db 1384 YYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYYVYY 1442

RESULT 2
US-09-252-991A-1304/c
; Sequence 1304, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1304
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1304

Query Match 3.5%; Score 50.2; DB 4; Length 1098;
Best Local Similarity 50.6%; Pred. No. 0.011;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 830 GGCTCCTCTCCACGTCGTCTGCTCTCCACGCGCGCCGCTCTACGAGGCGCTGGCACTGC 889
Db 602 GACTCCTGCTCAATGCTGCTGCGCGGCTGCGGAGCTGGTCTGGGGCGCGCTGATGGTGC 543
QY 890 TGACACCGGAGCCTTCGGCTCTTCGGGCTTCGGGCTTCGGCTTCATCTAGCGTGC 949
Db 542 TGGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 483
QY 950 CGTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTTCT 1009
Db 482 TGCTCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
QY 1010 GGTTCAGCTGGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
Db 422 TGGCGCTGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTC 364

RESULT 3
US-09-252-991A-1255

; Sequence 1255, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1255
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1255

Query Match 3.5%; Score 50.2; DB 4; Length 1257;
Best Local Similarity 50.6%; Pred. No. 0.011;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 830 GGCTCCTCTCCACGTCGTCTGCTCTCCACGCGCGCCGCTCTACGAGGCGCTGGCACTGC 889
Db 333 GACTCCTGCTCAATGCTGCTGCGCGGCTGCGGAGCTGGTCTGGGGCGCGCTGATGGTGC 392
QY 890 TGACACCGGAGCCTTCGGCTCTTCGGGCTTCGGGCTTCGGCTTCATCTAGCGTGC 949
Db 393 TGGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 452
QY 950 CGTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTTCT 1009
Db 453 TGCTCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
QY 1010 GGTTCAGCTGGCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
Db 513 TGGCGCTGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTC 571

RESULT 4
US-09-252-991A-1163
; Sequence 1163, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1163
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1163

Query Match 3.5%; Score 50.2; DB 4; Length 1746;
Best Local Similarity 50.6%; Pred. No. 0.013;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 830 GGCTCCTCTCCACGTCGTCTGCTCTCCACGCGCGCCGCTCTACGAGGCGCTGGCACTGC 889
Db 1283 GACTCCTGCTCAATGCTGCTGCGCGGCTGCGGAGCTGGTCTGGGGCGCGCTGATGGTGC 1342
QY 890 TGACACCGGAGCCTTCGGGCTTCGGGCTTCGGGCTTCGGCTTCATCTAGCGTGC 949

QY 1065 TCTCCAGTATGTTGGCCAGGCTCTTGGACCCCTTCTGGACCAAGCGCCAGGACTG 1124
 Db 946 CCGCGTCTCGTCTCCCGCTCTCGTCTCCACCTCCGGCCCGCAGCTCCTCCGTCACCT 1005
 QY 1125 CAGCCAGGAGAGAGGGGCTCACCTCTTATCTCGGCGACCCACTGCACAAGCAGGCGCG 1184
 Db 1006 CCGGCCAGCTCTCCATCACCTCCGGCCCGCAGCTCTCCATCACCTCCGGCGCGCCCA 1065
 QY 1185 TCTCCAGACTTAAATGTATCACCACTAAGCTGTGAGGGGACCCCAATCTGGACTCCTT 1244
 Db 1066 TCTCTGCACTTTCTCGGTCCCGCTCCCGCTCAGCACTTTCTCGGTCCCGCTCTCCAGCACCT 1125
 QY 1245 CCGCGCCTT 1253
 Db 1126 GCACCTCTT 1134

RESULT 13
 US-09-130-114-2
 ; Sequence 2, Application US/09130114
 ; Patent No. 5976807
 ; GENERAL INFORMATION:
 ; APPLICANT: Horlick, Robert A.
 ; APPLICANT: Damaj, Bassam B.
 ; APPLICANT: Robbins, Alan K.
 ; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
 ; TITLE OF INVENTION: From Multiple Transfected Episomes
 ; FILE REFERENCE: 0867/1D903US1
 ; CURRENT APPLICATION NUMBER: US/09/130,114
 ; CURRENT FILING DATE: 1998-08-06
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 1931
 ; TYPE: DNA
 ; ORGANISM: EBNA
 US-09-130-114-2

Query Match 3.1%; Score 45; DB 2; Length 1931;
 Best Local Similarity 44.1%; Pred. No. 0.22;
 Matches 189; Conservative 0; Mismatches 240; Indels 0; Gaps 0;
 QY 825 CTTCTGGCTCTCTCCACGCTGCTCTCCACGCGCGCCCGCTCTACGAGGCGCTGGC 884
 Db 706 CCGTCTCTCCCGCT 765
 QY 885 ACTGCTGACCAACCGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 944
 Db 766 CTTCTGGCT 825
 QY 945 CGTGGCT 1004
 Db 826 CTTCTGGCT 885
 QY 1005 CTTCTGGCT 1064
 Db 886 CTTCTGGCT 945
 QY 1065 TCTCCAGTATGTTGGCCAGGCTCTTGGACCCCTTCTGGACCAAGCGCCAGGACTG 1124
 Db 946 CCGCGTCTCGTCTCCCGCTCTCGTCTCCACCTCCGGCCCGCAGCTCTCCATCACCT 1005
 QY 1125 CAGCCAGGAGAGGGGCTCACCTCTTATCTCGGCGACCCACTGCACAAGCAGGCGCG 1184
 Db 1006 CCGGCCAGCTCTCCATCACCTCCGGCCCGCAGCTCTCCATCACCTCCGGCGCGCCCA 1065
 QY 1185 TCTCCAGACTTAAATGTATCACCACTAAGCTGTGAGGGGACCCCAATCTGGACTCCTT 1244
 Db 1066 TCTCTGCACTTTCTCGGTCCCGCTCCCGCTCAGCACTTTCTCGGTCCCGCTCTCCAGCACCT 1125
 QY 1245 CCGCGCCTT 1253

Db 1126 GCACCTCTT 1134
 RESULT 14
 US-09-485-549-1
 ; Sequence 1, Application US/09485549
 ; Patent No. 6361948
 ; GENERAL INFORMATION:
 ; APPLICANT: James Tricoli
 ; APPLICANT: Rachel Rhodinelli
 ; APPLICANT: Fox Chase Cancer Center
 ; TITLE OF INVENTION: Prognostic Compositions for Prostate Cancer and Methods of Use
 ; TITLE OF INVENTION: thereof
 ; FILE REFERENCE: FCCC 96-13
 ; CURRENT APPLICATION NUMBER: US/09/485,549
 ; CURRENT FILING DATE: 2000-11-09
 ; PRIOR APPLICATION NUMBER: PCT/US98/16768
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: 60/055,285
 ; PRIOR FILING DATE: 1997-08-13
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1835
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-485-549-1

Query Match 3.1%; Score 44.4; DB 4; Length 1835;
 Best Local Similarity 59.5%; Pred. No. 0.3;
 Matches 75; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 1316 CACTGAGCGCTGCTGGCGGAGGCTCTCGGACATCCGAGCAGGAGGAAAGTCTCCTGG 1375
 Db 1658 CACTTCCGGCCCTGGTGTGGGGGAGAGAGTGGGGTTCGGGGATCGGCAGTGGGAGG 1717
 QY 1376 GCGGATCTGTAATAAACCTTTTCTCTTTTCTTTTAAATAAAAAAAAAAAAAA 1435
 Db 1718 GCGGCTCTGAGATTAAAGAGTTTACCTTTGGGTAAAAATAAAAAAAAAAAAAA 1777
 QY 1436 AAAAAA 1441
 Db 1778 AAAAAA 1783

RESULT 15
 US-09-128-155-16
 ; Sequence 16, Application US/09128155
 ; Patent No. 6117654
 ; GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: 09404/052001
 ; CURRENT APPLICATION NUMBER: US/09/128,155
 ; CURRENT FILING DATE: 1998-08-03
 ; EARLIER APPLICATION NUMBER: US 60/091,650
 ; EARLIER FILING DATE: 1998-07-02
 ; EARLIER APPLICATION NUMBER: US 60/054,646
 ; EARLIER FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 16
 ; LENGTH: 152331
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(152331)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-128-155-16

Query Match 3.1%; Score 44.4; DB 3; Length 152331;

Best Local Similarity 51.5%; Pred. No. 1.2;
Matches 102; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy	832	CTCCTCTCCACGCTGCTGCTTCCACGCGGCCCCCGCTCTACGGAGGCGCTGGCCTGCTG 891
Db	22095	CCCGCCCCCCCCCCCCCGAGCGGACGCGCCCCCCCCCGCGCCCCCGCACCCCCG 22154
Qy	892	ACCACGGGAGCCTTCGGGCTCTTCGGGGTCTTCGGCTTGGCTCCATCTTAGCGTGCGG 951
Db	22155	ACCCCCCGCGCGCGCGCGCGCGCCCCCGCGCCCCCGCGCCCCCGCGCGCGCG 22214
Qy	952	CTCTGCGCGCTCCGCGCTAGGCTCCTCCGCGCTCACCCTCAGTACGGCGCGCTTCTGG 1011
Db	22215	CCCCACCCCCCCCCCGAGCGCGGACCGCGGCCCCCCCCCGACCCCCCGAGCCCCCGC 22274
Qy	1012	GTCACGCTGGCAACCGGC 1029
Db	22275	CCCCGCGCGCGACCGGC 22292

Search completed: June 3, 2004, 13:07:59
Job time : 127 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 10:08:49 ; Search time 653 Seconds
(without alignments)
10067.131 Million cell updates/sec

Title: US-09-936-456-1

Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2995936 seqs, 228098010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_NEW_PUB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1435.6	99.6	1474	16	US-10-264-237-1097
2	1413	98.1	1421	15	US-10-187-657-2
3	1407.8	97.7	1420	15	US-10-187-657-4
4	694.6	48.2	1594	15	US-10-187-657-9
5	451.6	31.3	522	15	US-10-187-657-5
6	396	27.5	2684	16	US-10-094-749-1195
7	385.8	26.8	450	15	US-10-187-657-7
8	354.8	24.6	1029	10	US-09-759-130B-424
9	354.8	24.6	1029	14	US-10-042-431-54
10	354.8	24.6	2133	10	US-09-759-130B-423
11	354.8	24.6	2133	14	US-10-042-431-53
12	322	22.3	506	15	US-10-187-657-6
13	300	20.8	346	15	US-10-187-657-3
14	245.8	17.1	2684	16	US-10-094-749-1195

15	200.6	13.9	439	13	US-10-027-632-91970	Sequence 91970, A
16	200.6	13.9	439	16	US-10-027-632-91970	Sequence 91970, A
17	196.2	13.6	232	9	US-09-783-590-9492	Sequence 9492, Ap
18	170	11.8	250	15	US-10-187-657-8	Sequence 8, Appli
19	64	4.4	512	15	US-10-106-698-3277	Sequence 3277, Ap
20	51.2	3.6	1123	13	US-10-425-114-15938	Sequence 15938, A
21	50.6	3.5	1455	15	US-10-156-761-5889	Sequence 5889, Ap
22	50.6	3.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
23	50	3.5	594	13	US-10-142-426-10	Sequence 10, Appl
24	50	3.5	594	15	US-10-123-155-10	Sequence 10, Appl
25	50	3.5	594	15	US-10-146-731-10	Sequence 10, Appl
26	50	3.5	594	15	US-10-140-472-10	Sequence 10, Appl
27	50	3.5	594	15	US-10-141-761-10	Sequence 10, Appl
28	50	3.5	594	15	US-10-142-885-10	Sequence 10, Appl
29	50	3.5	594	15	US-10-158-790-10	Sequence 10, Appl
30	50	3.5	594	16	US-10-137-871-10	Sequence 10, Appl
31	50	3.5	594	16	US-10-140-923-10	Sequence 10, Appl
32	50	3.5	594	16	US-10-141-756-10	Sequence 10, Appl
33	50	3.5	594	16	US-10-141-759-10	Sequence 10, Appl
34	50	3.5	594	16	US-10-140-803-10	Sequence 10, Appl
35	50	3.5	594	16	US-10-140-864-10	Sequence 10, Appl
36	50	3.5	1362	15	US-10-156-761-7248	Sequence 7248, Ap
37	50	3.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
38	49.4	3.4	2336	13	US-10-425-114-15797	Sequence 15797, A
39	48.8	3.4	919	13	US-10-425-114-22481	Sequence 22481, A
40	48.8	3.4	2109	13	US-10-425-114-32191	Sequence 32191, A
41	48.6	3.4	1262	13	US-10-425-114-23114	Sequence 23114, A
42	48.6	3.4	1339	13	US-10-425-114-33879	Sequence 33879, A
43	48.6	3.4	1352	13	US-10-425-114-13474	Sequence 13474, A
44	48.4	3.4	511	13	US-10-424-599-56520	Sequence 56520, A
45	48	3.3	234	10	US-09-814-353-4810	Sequence 4810, Ap

ALIGNMENTS

RESULT 1

US-10-264-237-1097
; Sequence 1097, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1097
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-1097

Query Match	99.6%	Score 1435.6;	DB 16;	Length 1474;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 1435;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	AAAGTAACGGCTACAGACAGTACAGAAATAGTTTCGCTCGCGCTAGAAAACTCTGTGC	60	
Db	27	AAAGTAACGGCTACAGACAGTACAGAAATAGTTTCGCTCGCGCTAGAAAACTCTGTGC	86	
Qy	61	GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAAACGGAGAGGTGCA	120	
Db	87	GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAAACGGAGAGGTGCA	146	
Qy	121	GGACTCAGACTTCACAGCCCACTCGGTCGCCGCTTGTACCAAGAGACGCCAAGGAC	180	
Db	147	GGACTCAGACTTCACAGCCCACTCGGTCGCCGCTTGTACCAAGAGACGCTCAAGGAC	206	


```

429 GTGAGATTCTTCTAGTCTGTTATAGGCGCAGAAATTTGGCTGTGCACTTCACTGCA 488
481 GAATGGTTGGTGGTACAGTGAACACCAACACATCTTACAAAGCCCTTCAGGCGAGCGGC 540
489 GAATGGTTGGTGGTACAGTGAACACCAACACATCTTACAAAGCCCTTCAGGCGAGCGGC 548
541 GTTACAGCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
549 GTTACAGCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 608
601 ACCCAGTGCATCAGCTGAACGAGACCATTTGACTACAAAGCAGATTCACTGGCGTCTG 660
609 ACCCAGTGCATCAGCTGAACGAGACCATTTGACTACAAAGCAGATTCACTGGCGTCTG 668
661 AAAGAGAAATTACGCGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGACCCAGTG 720
669 AAAGAGAAATTACGCGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGACCCAGTG 728
721 CTCTACCTGGCGGAGAAATTACACCGAGTAGCCCTTGGGCTGTACACAGGAGTACAC 780
729 CTCTACCTGGCGGAGAAATTACACCGAGTAGCCCTTGGGCTGTACACAGGAGTACAC 788
781 CTGGCGGAGACATACGCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGG 840
789 CTGGCGGAGACATACGCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGG 848
841 AACGTGCTGCTCTCCACCGCGCGGCTCTACGAGGAGGCTTGGGCTGTACACAGGAG 900
849 AACGTGCTGCTCTCCACCGCGCGGCTCTACGAGGAGGCTTGGGCTGTACACAGGAG 908
901 GCCTTGGCGCTCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 960
909 GCCTTGGCGCTCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 968
961 CTGGCGCTAGGCTCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1020
969 CTGGCGCTAGGCTCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1028
1021 GCAACCGGCTCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1080
1029 GCAACCGGCTCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1088
1081 CCCAGCGCTCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1140
1089 CCCAGCGCTCTTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1148
1141 GGCTCACCTCTTATCTCTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1200
1149 GGCTCACCTCTTATCTCTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1208
1201 TGATACACCACTTATCTCTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1260
1209 TGATACACCACTTATCTCTGGGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGG 1268
1261 CGCAGGCGGAGAGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGGAGAGAGAG 1320
1269 CGCAGGCGGAGAGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGGAGAGAGAG 1328
1321 AGCGCTGCTGGCGGAGAGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGGAGAGAG 1380
1329 AGCGCTGCTGGCGGAGAGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGCTGCGGAGAGAG 1388
1381 TCTGTAAATAAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1413
1389 TCTGTAAATAAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1421

```

RESULT 3

US-10-187-657-4

; Sequence 4, Application US/10187657

; Publication No. US20030068311A1

; GENERAL INFORMATION:

```

; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 3221661CA2
; US-10-187-657-4

```

Query Match 97.7%; Score 1407.8; DB 15; Length 1420;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1409; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGCTAGAAACCTCTGTGC 60
Db 9 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGCTAGAAACCTCTGTGC 68
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAAAGGAGAGGTGCA 120
Db 69 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTAAAGGAGAGGTGCA 128
QY 121 GGACTCAGACTTCACCGCCACTCGGTCCAGCCTTGTACGCAAGAGAGAGCCCAAGGAC 180
Db 129 GGACTCAGACTTCACCGCCACTCGGTCCAGCCTTGTACGCAAGAGAGAGCCCAAGGAC 188
QY 181 GGGCTCTCCGCGTCCAGGAGCGCCAGCTTGTGGCTTGCCTGCCGCTGCGTGCAGC 240
Db 189 GGGCTCTCCGCGTCCAGGAGCGCCAGCTTGTGGCTTGCCTGCCGCTGCGTGCAGC 248
QY 241 ACTGGCGCGGCTGAGCATGACATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCGAGCC 300
Db 249 ACTGGCGCGGCTGAGCATGACATGACCCCTGTGGAACGGCGTACTGCTTTTACCCCGAGCC 308
QY 301 CGGCATGCCGAGGCTTCAGCGTTCACGCTTCCACTGCTCATGTTATTTAGTGTGTTGGCTCTA 360
Db 309 CGGCATGCCGAGGCTTCAGCGTTCACGCTTCCACTGCTCATGTTATTTAGTGTGTTGGCTCTA 368
QY 361 GCAGCAAGCTTCTGCTCATGTTTCCGCGGATCCGTGGCCACTCGCGCTGTTGGTTG 420
Db 369 GCAGCAAGCTTCTGCTCATGTTTCCGCGGATCCGTGGCCACTCGCGCTGTTGGTTG 428
QY 421 GTGAGAGTTCTTCTCAGTCTGTTTCATAGGCGCAGAAATTTGGCTGTGCACTTCAGTGCA 480
Db 429 GTGAGAGTTCTTCTCAGTCTGTTTCATAGGCGCAGAAATTTGGCTGTGCACTTCAGTGCA 488
QY 481 GAATGGTTCTGGGTACAGTGAACACCAACACATCTTACAAAGCCCTTCAGGCGAGCGGC 540
Db 489 GAATGGTTCTGGGTACAGTGAACACCAACACATCTTACAAAGCCCTTCAGGCGAGCGGC 548
QY 541 GTTACAGCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
Db 549 GTTACAGCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 608
QY 601 ACCCAGTGCATCAGCTGAACGAGACCATTTGACTACAAAGCAGATTCACTGGCGTCTG 660
Db 609 ACCCAGTGCATCAGCTGAACGAGACCATTTGACTACAAAGCAGATTCACTGGCGTCTG 668
QY 661 AAAGAGAAATTACGCGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGACCCAGTG 720
Db 669 AAAGAGAAATTACGCGCGGAGTACGCGAAGCAGTGGAGAGAGGGGCTGCGGACCCAGTG 728

```

Query Match	48.2%;	Score 694.6;	DB 15;	Length 1594;
Best Local Similarity	75.7%;	Pred. No. 1.8e-185;		
Matches 918;	Conservative 0;	Mismatches 279;	Indels 16;	Gaps 4;
Qy 222	CTGCCGCTCGGTGAGCAGCTCGGCGCGGTGAGCATGACCCCTGTGGACGGCGTACT	281		
Db 107	CTCGCTGCCTGCGAACAGTCTCGAACTCTTGCAAGATGACTGCGTGGGACGGGTGCT	166		
Qy 282	GCCTTTTACCCCGAGCCCGGCATGCCCGCAGGCTTCAGCGTTCCACTGCTCATCGTTAT	341		
Db 167	ACCCCTTTACCCCGAGCCCGGCATGCCCGCAGGCTTCAGCGTACCGTCTCATTTGTTAT	226		
Qy 342	TCTAGTGTGTTTGGTCTTAGCAGCAAGCTTCTTGCTCATCTTGCCCGGGATCCGTGGCCA	401		
Db 227	CCTGGTGTCTTGTCTTTGGCTGCTAGCTTCTGTGTCATCTTGCCCTGGAATCCGTGGCA	286		
Qy 402	CTCGCGCTGTTTGGTGGTGAGAGTTCCTCTCAGTCTGTTCTCATAGCGCGCAGAAATGT	461		
Db 287	CTCGCGCTGTTCTGGTGGTGAGAGTCTCTTTAGCTGTTCTCATAGGTGCCGAATGT	346		
Qy 462	GGCTGTGCACCTTCAGTGCAGAAATGGTTCTGTTGGTACAGTGAACACCAACACATCCTACAA	521		
Db 347	GGCTGTGCACCTTCAGTGCAGAAATGGTTCTGTTGGTACAGTGAACACCAACACATCCTACAA	406		
Qy 522	AGCCTTCAGCGCAGCGCGTTACAGCGCGTGTGCGTCTGCTCGTGGGCTTGAGGGCAT	581		
Db 407	AGCCTTCAGTCCATCCCGTGTTCAGTCCATGTGCGTCTGACGCTGGGCTTGCGGGCT	466		
Qy 582	TAATATTACACTCACAGGAGACCCAGTGCATCAGCTGAACGAGACCATTTGACTACAAACA	641		
Db 467	TAAATATTACACTCCGAGGAACACCGAGGAGCAGCTGAACGAGACCATTTGACTACAAACA	526		
Qy 642	GCAGTTTCACTTGGCGTCTGAAAGAGAAATACCGCGGAGTACCGGAACGCACTGGAGAA	701		
Db 527	GGCTTTTCACTTGGCGTCTGAAAGAGAAATACCGCGGAGTACCGGAACGCACTGGAGAA	586		
Qy 702	GGGCTGCCGGAACCACTGCTCTACTCTGGCGGAGAAATTCACACCGAGTAGCCCTTGGCG	761		
Db 587	GGGCTGCCGGAACCACTGCTCTACTCTGGCGGAGAAATTCACACCGAGTAGCCCTTGGCG	646		
Qy 762	CCTGTATCCACAGTACCACTGGCGGACACTACCGCTCGGCCACCGCTATGGGTGGCGTT	821		
Db 647	GCTGTATCCACCAATATCACCTGGCGGCTCACTATGGCGGAGCAACTGTGGGTGGCAAT	706		

Db	827	CAGCGTGC	CGCTCTG	CCACTTC	CGCCTCG	GGCTC	CGCGGT	CCTC	ADGCTT	ACTACG	CGC	886
Qy	1002	CGCCTTCT	GGGTAC	CGCTGG	CAACCG	CGCGT	CCTGTG	CTCT	CTTCG	GAGGG	CGCGT	1061
Db	897	CTCCTTTT	GGCTAC	CGTGGC	CAACCG	CGCAT	CCTGAG	CTCT	CTCG	AGGG	CGCGT	946
Qy	1062	GATCTCC	AGTAGT	TCGGG	CCCAG	CGCTCT	TCGAC	CCCTT	CTGG	ACCA	AGCG	1121
Db	947	GATTCTC	CACTAT	ACTCGG	CCCAG	CGCCCT	TCGCT	CTCT	CTGG	ATCT	AGTCT	1006
Qy	1122	CT---	GCAGCC	AGGAG	AGGGG	GCTCAC	CTCT	TTAT	CTCT	GGCG	ACCC	1178
Db	1007	CTGTAG	CAAC	CAGGCT	AAAGG	AACT	CAC	CTCT	CTC	ACCT	CAAC	1066
Qy	1179	GGCGCT	CTCC	CAGACT	TAAAT	GTAT	CAC	CACT	TAAC	CTGT	GAGGG	1238

Db 1067 GTTGAAGAGTCCAGACTTAAA--TATTACCACTCTCTCTGTGAAATAAATACTGACTCCGGA 1123
QY 1239 CTCCTTCCCGCCTTGGACATGCGAGCGCGGGAAGCAAGTCCCGCCAGGCCTG-GGCCA 1297
Db 1124 TTTCTACCCCTCTTTGGGACCCCATAGACCTGAAGACTGTGTTAAGAGCGCGTGCCAGGA 1183
QY 1298 GGAGAGCTCCAGGAAGGGCACTGAGCGCTGCTGGCGGAGGCCCTCGGACATCCGCAAGCA 1357
Db 1184 GCATAGCTTAGTCTGGAGAACTGCTTGTCCCTGCGGGAGCAACAAGGGGGCACCCCATGTG 1243
QY 1358 CC-----AGGGAAGTCTCCTGGGCGCATCTGTAAATAAACCTTTTCTTTTGT 1408
Db 1244 CCTTGTCTAGAGAAATGATTTCTCAGGAGAACTGTAAATAAACTTTTGTCTTTT 1303
QY 1409 TTTTAAATAAAAA 1421
Db 1304 TTTTCTTCAAAA 1316

RESULT 5
US-10-187-657-5
; Sequence 5, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 3221661R6
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 455, 480, 483
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-5

Query Match 31.3%; Score 451.6; DB 15; Length 522;
Best Local Similarity 96.7%; Pred. No. 5e-117;
Matches 502; Conservative 0; Mismatches 12; Indels 5; Gaps 4;
QY 1 AAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTG 60
Db 9 AAGTAACGGCTACAGACAGTGAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTG 68
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACAGCTTCCTTAACGGAGGTGCA 120
Db 69 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACAGCTTCCTTAACGGAGGTGCA 128
QY 121 GGACTCAGACTTCACAGCCCACTCGGTCCCAGCCTTGACGCAAGAGACGCCAAGGAC 180
Db 129 GGACTCAGACTTCACAGCCCACTCGGTCCCAGCCTTGACGCAAGAGAGACGCCAAGGAC 188
QY 181 GCGCTCTCCCGGTCCAGGCAGCCCGCAGCTTGCTGGCTTGCCTGCCCGCTGCGTGCAGC 240
Db 189 GCGCTCTCCCGGTCCAGGCAGCCCGCAGCTTGCTGGCTTGCCTGCCCGCTGCGTGCAGC 248
QY 241 ACTCGCGGGCGTGACGATGACCTGTGGAAACGGCGTACTGCTTTTACCCCGAGCCC 300

Db 249 ACTCGGCGCGGTGCAGCATGACCCCTGTGGAAACGGCGTACTGCTTTTACCCCGAGCCC 308
QY 301 CGGCATGCCCGCAGGCTTCAGCGTTCCACTGCTCATCGTTATTCTAGTGTTTTGGCTCTA 360
Db 309 CGGCATGCCCGCAGGTTTCAAGCGTTCCACTGCTCATCGTTATTCTAGTGTTTTGG-TCTA 367
QY 361 GCAGCAAGCTTCTCTGCTCATCTTGC CGGGGATCGGTGGCCACTCGCGCTGGTTTGGTTG 420
Db 368 GCAGCAAG--TTCTGTCTATCTTGC CGGGGATCGGTGGCCACTCGCGCTGGTTTGGTTG 425
QY 421 GTGAGAGTTCTTCTCAGTCTGTTCATAGGCGCAGAAATTGTGCTGTGCACTTCAAGTGCA 480
Db 426 GTGAGAGTTCTTCTCAGTCTGTTCATAGGCGCAGAAATTGTGCTGTGCACTTCTCNGT-NA 484
QY 481 GAATGGTTCTGTTGAGTACAGTGAACACCAACACATCCTAC 519
Db 485 GAATGGTTCTGTTGAGTACATTG-ACACCAACACATCCTAC 522

RESULT 6
US-10-094-749-1195/c
; Sequence 1195, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1195
; LENGTH: 2684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-1195

Query Match 27.5%; Score 396; DB 16; Length 2684;
Best Local Similarity 97.6%; Pred. No. 5.5e-101;
Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1007 TCTGGGTCAAGCTGGCAACCGCGCTCCTGTGCTCTTCTCGAGGGGCGGTGAGTC 1066
Db 2363 TTTCGATCCCAACCGCCACAGGCGTCTGTGCTCTTCTCGAGGGGCGGTGAGTC 2304
QY 1067 TCCAGTATGTTGGGCGGCGCTCTTCGCAACCTTCTTGACAAAGCGCCAGGACTGCA 1126
Db 2303 TCCAGTATGTTGGGCGGCGCTCTTCGCAACCTTCTTGACAAAGCGCCAGGACTGCA 2244
QY 1127 GCCAGGAGAGAGGGGGCTCACTCTTATCTCGGCGACCCACTGTCACAAGAGGCGGCTC 1186
Db 2243 GCCAGGAGAGAGGGGGCTCACTCTTATCTCGGCGACCCACTGTCACAAGAGGCGGCTC 2184

QY	1187	TCCAGACTTAAATGTATCACCACTAAACCTGTGAGGGGGACCCCAATCTGGACTCTCTCC	1246
Db	2183	TCCAGACTTAAATGTATCACCACTAAACCTGTGAGGGGGACCCCAATCTGCACTCTCTCC	2124
QY	1247	COGCTTTGGGACATCGCAGGCGCGGGAAGCAGTGCCCGCCAGGCCCTGGGCCCAGGAGAGCTC	1306
Db	2123	COGCTTTGGGACATCGCAGGCGCGGGAAGCAGTGCCCGCCAGGCCCTGGGCCCAGGAGAGCTC	2064
QY	1307	CAGGAAGGGCACTGAGCGCTGCTGGCGGAGGCGCTGGACATCCGCAGGCACCCAGGGAAA	1366
Db	2063	CAGGAAGGGCACTGAGCGCTGCTGGCGGAGGCGCTGGACATCCGCAGGCACCCAGGGAAA	2004
QY	1367	GTCTCTGGGGCGATCTGTAAATAAACCTTTTTTTCTTTTGTTTTTTAAAAA	1418
Db	2003	GTCTCTGGGGCGATCTGTAAATAAACCTTTTTTTCTTTTGTTTTTTAAAAA	1952

RESULT: 7

```

US-10-187-657-7
; sequence 7, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 SBQA04510D1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 140, 165, 203, 237, 269, 289, 307, 338, 344, 347, 354, 390, 430, 4
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-7

```

```
Query Match      26.8%; Score 385.8; DB 15; Length 450;
Best Local Similarity 93.7%; Pred.No. 1.8e-98;
Matches 404; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
```

Qy	616	CTGAACGAGACCAATTGACTACAA	CGAGCAGTTCACCTGGCGTCTGAAAGAGAAATTACGCC	675
Db	21	CTGAACGAGACCAATTGACTACAA	CGAGCAGTTCACCTGGCGTCTGAAAGAGAAATTACGCC	80
Qy	676	GCGGAGTACGGGAACGCACTGGAGAGGGGCTGCCGACCCAGTGCTCTACCTGGCGGAG	735	
Db	81	GCGGAGTACGGGAACGCACTGGAGAGGGGCTGCCGACCCAGTGCTCTACCTGGCGGAN	140	
Qy	736	AAGTTCACACCGAGTAGCCCTTGCGGCCCTGTACCA	CCAGTACCACTGGCGGGACACTAC	795
Db	141	AAGTTCACACCGAGTAGCCCTTGCGGCCCTGTACCA	CCAGTACCACTGGCGGGACACTAC	200
Qy	796	GCTCGGCGCACGCTATGGGTGGCGTTCTGCTTCTGGCTCCTCTCCAAAGTGTGCTCTCC	855	
Db	201	GCTCGGCGCACGCTATGGGTGGCGTTCTGCTTCTGGNTCCTCTCCAAAGTGTGCTCTCC	260	
Qy	856	ACGCGCGGCCCGCTCTACGGAGGCCCTGGCACTGCTGACCA	CCGGAGCCTTGGCGCTCTTC	915
Db	261	ACGCGCGGCCCGCTCTACGGAGGCCCTGGNACTGCTGACACCGGAGNCTTGGCGCTCTTC	320	

US-10-187-657-6

Query Match 22.3%; Score 322; DB 15; Length 506;
Best Local Similarity 86.3%; Pred. No. 2e-80;
Matches 353; Conservative 0; Mismatches 54; Indels 2; Gaps 2;

QY 464 CTGTGCACTTCAGTGCAGAAATGTTCTGTTGAGTACAGTGAACACCAACATCTTACAAAG 523
DB 17 CTGTGCACTTCAGTGCAGAAATGTTCTGTTGAGTACAGTGAACACCAACATCTTACAAAG 75

QY 524 CTTTACGCGCAGCGCGTTTACAGCCCGTGTCTGTTGAGTACAGTGAACACCAACATCTTACAAAG 583
DB 76 CTTTACGCGCAGCGCGTTTACAGCCCGTGTCTGTTGAGTACAGTGAACACCAACATCTTACAAAG 135

QY 584 ATATTACATCTACAGGGACCCAGTGCATCAGTGAACAGAGACCAATGACTACACGAGC 643
DB 136 ATATTACATCTACAGGGACCCAGTGCATCAGTGAACAGAGACCAATGACTACACGAGC 195

QY 644 AGTTACCTGGCGTCTGAAAGAGATTACGCCCGGAGTACCGGACGACCTGGAGAGG 703
DB 196 AGTTACCTGGCGTCTGAAAGAGATTACGCCCGGAGTACCGGACGACCTGGAGAGG 255

QY 704 GGCTCCGGACCCAGTGTCTTACTTGGCGGAGAGTTTACACCGAGTAGCCCTTGGGCGC 763
DB 256 GGCTCCGGACCCAGTGTCTTACTTGGCGGAGAGTTTACACCGAGTAGCCCTTGGGCGC 315

QY 764 TGTACCAACAGTACACCTGGCGGACACTACGCTTGGCGGACGCTATGGTGGCGTCT 823
DB 316 TGTNNANCNAGNACCAANTGCGGACNNATACANTCGGCCACGCTATGGTGGCGTCT 375

QY 824 GCTTCTGGCTCC-TCTCCAACGTTGCTCTCTCCACGCGCGGCCCGCTCT 871
DB 376 GNTTCTGGCTCTTNTTCCAGGTGCTGTGTTCTNNANGNGGCCNCGT 424

RESULT 13
US-10-187-657-6
; Sequence 3, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9

US-10-187-657-6

Query Match 22.3%; Score 322; DB 15; Length 506;
Best Local Similarity 86.3%; Pred. No. 2e-80;
Matches 353; Conservative 0; Mismatches 54; Indels 2; Gaps 2;

QY 464 CTGTGCACTTCAGTGCAGAAATGTTCTGTTGAGTACAGTGAACACCAACATCTTACAAAG 523
DB 17 CTGTGCACTTCAGTGCAGAAATGTTCTGTTGAGTACAGTGAACACCAACATCTTACAAAG 75

QY 524 CTTTACGCGCAGCGCGTTTACAGCCCGTGTCTGTTGAGTACAGTGAACACCAACATCTTACAAAG 583
DB 76 CTTTACGCGCAGCGCGTTTACAGCCCGTGTCTGTTGAGTACAGTGAACACCAACATCTTACAAAG 135

QY 584 ATATTACATCTACAGGGACCCAGTGCATCAGTGAACAGAGACCAATGACTACACGAGC 643
DB 136 ATATTACATCTACAGGGACCCAGTGCATCAGTGAACAGAGACCAATGACTACACGAGC 195

QY 644 AGTTACCTGGCGTCTGAAAGAGATTACGCCCGGAGTACCGGACGACCTGGAGAGG 703
DB 196 AGTTACCTGGCGTCTGAAAGAGATTACGCCCGGAGTACCGGACGACCTGGAGAGG 255

QY 704 GGCTCCGGACCCAGTGTCTTACTTGGCGGAGAGTTTACACCGAGTAGCCCTTGGGCGC 763
DB 256 GGCTCCGGACCCAGTGTCTTACTTGGCGGAGAGTTTACACCGAGTAGCCCTTGGGCGC 315

QY 764 TGTACCAACAGTACACCTGGCGGACACTACGCTTGGCGGACGCTATGGTGGCGTCT 823
DB 316 TGTNNANCNAGNACCAANTGCGGACNNATACANTCGGCCACGCTATGGTGGCGTCT 375

QY 824 GCTTCTGGCTCC-TCTCCAACGTTGCTCTCTCCACGCGCGGCCCGCTCT 871
DB 376 GNTTCTGGCTCTTNTTCCAGGTGCTGTGTTCTNNANGNGGCCNCGT 424

RESULT 12
US-10-187-657-6
; Sequence 6, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657

QY 255 CAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCGGCGATGCGCAGG 314
DB 66 CAAGATGGCTACTTTGGGACACACATTCCTTCTATGCTGGCCCAAGCCAACTTCCC 125

QY 315 CTTTACGCTTCCACTGCTCATCGTTTATCTAGTGTGTTTGGCTTACGAGCAAGCTTCT 374
DB 126 GATGACACCACTTTTGGCCAGCATCATCATGATCTTCTGACTGCACTGGCCACGTTCA 185

QY 375 GCTCATCTTGGCGGATCCGTGGCCACTGCGTGGTGTGTTGGTGGAGATTTCTTCT 434
DB 186 CGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 245

QY 435 CAGTCTGTTTATAGGGCAGAAATGTTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
DB 246 CAGCTTATTCATCGGGCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305

QY 495 TACAGTGAACACCAACACATCTTACAAAGCCTTACAGCGCAGCGCGTTACAGCCCGTGT 554
DB 306 CCAGTTCAGCACCAACACATCATACAAGGCCTTCAAGGCCTTCAAGGCCTTCAAGGCCT 365

QY 555 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614
DB 366 TGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425

QY 615 GCTGAACGAGACCTTGAATCAACAGGAGTTCACCTGGCTGCTGCTGCTGCTGCTGCTGCT 674
DB 426 GCTGAATGAGACCATCAATTACACAGGAGTTCACCTGGCTGCTGCTGCTGCTGCTGCTGCT 485

QY 675 CGCGGAGTACGGAACGCACTGGAGAGGGGCTGCGGACCCAGTGTCTTACCTGGCGGA 734
DB 486 TGAGGAGTGTCAAGGCTCTGGAGAGGGGCTGCGGACCCCTGCTGCTGCTGCTGCTGCTGCT 545

QY 735 GAAGTTACACCGAGTACCTTGGCGCTGTACACAGGAGTTCACCTGGCGGAGCACTA 794
DB 546 GAAGTTACCTCAAGAGCCCATGTGCTGCTTACCGGAGTACCGGAGTACCGGAGTACCGGAG 605

QY 795 CGCCTCGGCGCAGCTATGGGTGGGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854
DB 606 CACCTCAGGCACTGCTATGGGTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665

QY 855 CACGCGGCGCGCTCTACGAGGCGCTGCGCACTGCTGACACCGAGGCTTCCGCTGCTT 914
DB 666 CATGCTGCTGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 725

QY 915 CG--GGGCTTGGCTTGGCTTCCATCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
DB 726 GGCTCTGCTTCTTCTCCATGGCCACATCACTACCTCACCTGCTGCTGCTGCTGCTGCTGCTGCT 785

QY 972 CTCCTCGGCGCTCACCACCTCAGTACGCGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
DB 786 CGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845

QY 1032 CTTGCTGCTTCTTCTGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
DB 846 GCT 905

QY 1092 TCGCACCTTCTGACCAAGGCGCAAGGA 1121
DB 906 GAAGGCTTCTTCAACACGAGTGTGGATGA 935

RESULT 12
US-10-187-657-6
; Sequence 6, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657

SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 1752794F6
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 91, 188, 206, 291, 337
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-3

Query Match 20.8%; Score 300; DB 15; Length 346;
Best Local Similarity 94.6%; Pred. No. 2.8e-74;
Matches 330; Conservative 0; Mismatches 15; Indels 4; Gaps 2;
QY 927 CTTGGCCTCCATCTCTAGGTCGCGCTCTGCGGCTCGCGCTAGGCTCCTCGGCTCAC 986
Db 1 CTTGGCCTCCATCTCTAGGTCGCGCTCTGCGGCTCGCGCTAGGCTCCTCGGCTCAC 60
QY 987 CACTCAGTACGGCGCGCTCTGCGGCTCGCGCTAGGTCGCGCTAGGCTCCTCGGCTCAC 1046
Db 61 CACTCAGTACGGCGCGCTCTGCGGCTCGCGCTAGGTCGCGCTAGGCTCCTCGGCTCAC 120
QY 1047 CGGAGGGCGGCTGAGTCTCCAGTATGTTCCGCGGCTCGCGCTAGGTCGCGCTAGGTCGCGG 1106
Db 121 CGGAGGGCGGCTGAGTCTCCAGTATGTTCCGCGGCTCGCGCTAGGTCGCGCTAGGTCGCGG 180
QY 1107 CCAAAGCGCAAGGACTGCGGCTCGCGCTAGGTCGCGCTAGGTCGCGCTAGGTCGCGG 1165
Db 181 CCAAAGCGCAAGGACTGCGGCTCGCGCTAGGTCGCGCTAGGTCGCGCTAGGTCGCGG 240
QY 1166 CACTGCAAGCAGGCGGCTCTCCAGTATGTTCCGCGGCTCGCGCTAGGTCGCGCTAGGTCGCGG 1225
Db 241 CACTGCAAGCAGGCGGCTCTCCAGTATGTTCCGCGGCTCGCGCTAGGTCGCGCTAGGTCGCGG 300
QY 1226 GACCCAAATCTGGACTCTTCCCGGCTCGCGCTAGGTCGCGCTAGGTCGCGCTAGGTCGCGG 1274
Db 301 GACCCAAATCTGGACTCTT---CCCGGCTCGCGCTAGGTCGCGCTAGGTCGCGCTAGGTCGCGG 346

RESULT 14
US-10-094-749-1195
; Sequence 1195, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381

PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1195
; LENGTH: 2684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-1195
Query Match 17.1%; Score 245.8; DB 16; Length 2684;
Best Local Similarity 67.4%; Pred. No. 1.5e-58;
Matches 362; Conservative 0; Mismatches 172; Indels 3; Gaps 1;
QY 588 TACACTCACAGGAGCCAGTGCATCAGCTGAACGAGACCATTTGACTACAAACGAGCAGTT 647
Db 250 TGCAATCTCTGGGACCCCGCTGCGAGCAGCTGAATGAGACCATCAATTACAACGAGGAGTT 309
QY 648 CACCTGGCGTCTGAAAGAGAAATTAGCCGCGGAGTACGGAACGCACTGGAGAGGGGCT 707
Db 310 CACCTGGCGCTGGTGAAGAACTATGCTGAGGAGTATGCAAGGCTCTGGAGAGGGGCT 369
QY 708 GCGGACCCAGTGTCTACTCTGGGAGAGAAATTACACCGAGTAGCCCTTTCGGGCTGTA 767
Db 370 GCGGACCCCTGTGTGTACCTAGTGAAGAGTTCACTCAAGAGCCCATGTGGGCTATA 429
QY 768 CCACAGTACCACTGGGCGGAGACTACGCTCGGCGGCGGCTTACGAGGCGCTTCTGCTT 827
Db 430 CGGCAAGTACCGCTGGGCGGAGACTACACCTCAGCCATGCTATGGTGGCATTCCTCTG 489
QY 828 CTGGCTCTCTCCAAAGTGTCTGCTTCCAGCGGCGGCGGCTTACGAGGCGCTTGGCACT 887
Db 490 CTGGCTCTGCGCAATGTATGCTCTCCAGCTGTGGCTCTCTCTCTCTCTCTCTCTCTCT 549
QY 888 GCTGACCAACGAGGAGCTTCCGCTCTTCG--CGGCTCTTGGCTTGGCTTGGCTTGGCTT 944
Db 550 ATTGGCCACCGGCAATGTATGCTCTCCAGCTGTGGCTCTCTCTCTCTCTCTCTCTCTCT 609
QY 945 CGTGGCGCTCTGCGGCTTCCGCTTGGCTTCCAGCGGCGGCTTCCAGCTCAGTACGCGCGC 1004
Db 610 CACCTCACCTGTCCCTTCCAGCTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 669
QY 1005 CTTCTGGTCAACGCTGGCAACCGGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1064
Db 670 CTTCTGGTCAACGCTGGCAACCGGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 729
QY 1065 TCTCCAGTATGTTCCGCGGCGGCTTCTGCGGCGGCTTCTGCGGCGGCGGCGGCGG 1121
Db 730 GGCCACAGGATGCGGCGGCTTCAAGGCTTCTTCAACAGAGTGTGGATGA 786
RESULT 15
US-10-027-632-91970
; Sequence 91970, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91970
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (439)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-91970

Query Match	13.9%	Score 200.6;	DB 13;	Length 439;
Best Local Similarity	97.3%	Pred. No. 3.8e-46;		
Matches	214;	Conservative 0;	Mismatches 5;	Indels 1; Gaps 1;

Qy	812	GGGTGGCGTTCTGCTTCTGGCTCCTCTCCAAAGTGTGCTCTCCACGGCGGCCCGCTCT	871
Db	133	GGGTGGCGTTCTGCTTCTGGCTCCTCTCCAAAGTGTGCTCTCCACGGCGGCCCGCTCT	192
Qy	872	ACGGAGGCTGGCAGTGTG-ACCACGGAGCCTTCGGGCTCTTCGGGGTCTTCGCCCTTG	930
Db	193	GCGGAGGCTGGCAGTGTGCTENACCAAGGAGCCTTCGGGCTCTTCGGGGTCTTCGCCCTTG	252
Qy	931	GCCTCCATCTCTAGGCTGCGCTCTGCGGCTCCTCGGCTAGGCTCCTCGGCTCACCCT	990
Db	253	GCCTCCATCTCTAGGCTGCGGCTCTGCGGCTCCTCGGCTAGGCTCCTCGGCTCACCCT	312
Qy	991	CAGTACGGCGCGCCTTCTGGGTACGCTGGCAACGGCG	1030
Db	313	CAGTACGGCGCGCCTTCTGGGTACGCTGGCGACGGTG	352

Search completed: June 3, 2004, 13:19:21
Job time : 672 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 09:02:12 ; Search time 4061 Seconds
(without alignments)
10596.258 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	694.6	48.2	1594	11 AK018569	AK018569 Mus muscu
2	657	45.6	1309	11 AK008816	AK008816 Mus muscu
3	529.6	36.8	537	9 AI821606	AI821606 nk08all.x
4	515	35.7	515	9 AI826629	AI826629 wk35e04.x

C	5	510	35.4	510	9	AI304327	qo57h06.x
C	6	489.2	33.9	505	9	AI791844	nk08all.y
C	7	482.4	33.5	486	9	AI660560	we8b08.x
C	8	466	32.3	1014	13	BY708714	BY708714
C	9	446	31.0	446	9	AI984141	wu21c02.x
C	10	445.4	30.9	466	9	AA573825	nk08all.s
C	11	434	30.1	434	9	AI991272	wu41h04.x
C	12	414	28.7	414	9	AA593860	nn19f03.s
C	13	409.4	28.4	411	9	AI983793	wu20c09.x
C	14	405	28.1	405	9	AI274929	ql49c11.x
C	15	400.2	27.8	405	9	AI346155	qp43f12.x
C	16	392	27.2	538	9	AI660493	we67h02.x
C	17	388.4	27.0	538	13	BX517021	BX517021
C	18	385	26.7	573	14	CB854121	UI-CF-DU1
C	19	381	26.4	381	9	AI281211	qk5e08.x
C	20	375.8	25.1	737	14	CB305399	UI-CF-EN1
C	21	373.6	25.9	660	14	CF766127	CES004870
C	22	372.8	25.9	646	13	BY720747	BY720747
C	23	367.4	25.5	713	14	CB171450	CB171450 JID602700
C	24	356.4	24.7	1032	29	AY418362	AY418362 Homo sapi
C	25	349.8	24.3	1026	29	AY418364	Mus muscu
C	26	346	24.0	483	9	AI019726	AI019726 ua94c11.x
C	27	335	23.2	348	9	AW050605	AW050605 wz19b11.x
C	28	333.4	23.1	690	12	BM977010	UI-CF-EN1
C	29	333.4	23.1	896	12	BG968477	BG968477 602835393
C	30	324	22.5	687	14	CB850544	CB850544 UI-CF-EN1
C	31	322.4	22.4	324	12	BM987789	BM987789 UI-H-COO-
C	32	308.8	21.4	871	12	B1526485	B1526485 602925070
C	33	304.8	21.2	919	12	B1412602	B1412602 602990526
C	34	296	20.5	296	9	AI821178	AI821178 nel7a03.y
C	35	295.2	20.5	999	12	BI410133	BI410133 602964073
C	36	290.6	20.2	315	9	AA618335	AA618335 nq15g11.s
C	37	289.4	20.1	895	12	BI413312	BI413312 602986494
C	38	277	19.2	277	9	AI732165	AI732165 nel7a03.x
C	39	275	19.1	608	14	CF108325	CF108325 Shultzomi
C	40	274.4	19.0	654	10	BE914798	BE914798 601667822
C	41	263.2	18.3	766	29	AY418363	AY418363 Pan trogl
C	42	260	18.0	260	9	AW009962	AW009962 ws89h06.x
C	43	257.2	17.8	675	12	B1154672	B1154672 602904653
C	44	252	17.5	252	9	AI262416	AI262416 qk38e04.x
C	45	251.8	17.5	1040	12	BI554670	BI554670 603236624

ALIGNMENTS

RESULT 1	AK018569	1594 bp	mRNA	linear	HTC 20-SEP-2003
LOCUS	Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030623N16				
DEFINITION	sequence.				
ACCESSION	AK018569				
VERSION	AK018569.1	GI:12858338			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				

QY	1002	CGCCTCTGGTCACTGGCAACCGCGTCTCTGCTCTCTCTCGGAGGGCGGTGGT	1061
Db	887	CTCCCTTGGCTACGCTGGCCACCGCATCTCTGAGCTCTCTCTCGGAGGGCGGTGGT	946
QY	1062	GAGTCTCCAGTATGTTGGCCCGCAGCGCTCTTCGACCTCTCTGACCAAAAGCGCAAGGA	1121
Db	947	GATTCCTCACTATACTCGGCCCGCAGCGCTCTCTGCTCTCTCTGATCTAAGTGTCAAGA	1006
QY	1122	CT---GCAGCCAGGAGAGAGGGGGCTCACCTCTTATCTCGGAGCCCACTGCACAAGCA	1178
Db	1007	CTGTAGCAACAGGCTAAAGGAACTCACCTCTCACCTCAACAACCGCAACGAACA	1066
QY	1179	GGCGGCTCTCCAGACTTAAATGTATCACCACTAAGCTGTGAGGGGACCCAACTCGGA	1238
Db	1067	GTTGAAGAGTCCAGACTTAAA---TATTAACACTCTCTCTGTGAAATAACTGACTCCGA	1123
QY	1239	CTCCTTCCCGCTTGGACATCGAGCGCGGAGAGCTGCGCGGAGGCGCTG-GGCCA	1297
Db	1124	TTTCTACCCCTCTTTGGGACCCCATAGACCTGAAGACTGTGTTAAGAGCGCGTGGCAGGA	1183
QY	1298	GGAGAGTCCAGGAAGGCACTGAGCGCTGCTGGCGGAGGCTCGGACATCCGACGCA	1357
Db	1184	GCATAGCTTAGTCTGGAGAACTGCTTGTCTCTCGGAGCAGCAAGGGGCAACCCATGTG	1243
QY	1358	CC-----AGGAAAGTCTCTCTGGGCGATCTGTAATAAACCTTTTCTTTTGT	1408
Db	1244	CCTTGTCTAGAGAAATGATTTCTCAGGAGAACTGTAATAAACCTTTTCTTTTGT	1303
QY	1409	TTTTTAAAAAAA 1421	
Db	1304	TTTTTCTCAAAA 1316	
RESULT 2			
AK008816			
LOCUS		1309 bp mRNA linear HTC 20-SEP-2003	
DEFINITION		Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210403N03 product:hypothetical protein, full insert sequence.	
ACCESSION		AK008816	
VERSION		AK008816.1 GI:12843238	
KEYWORDS		HTC; CAP trapper.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		Carninci, P. and Hayashizaki, Y.	
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE		99279253	
PUBMED		10349636	
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	
MEDLINE		20493374	
PUBMED		11042159	
AUTHORS		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE		20530913	

11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

6 (bases 1 to 1309)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Iizawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

Location/Qualifiers

1. .1309

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:2210403N03"

/db_xref="MGI:1893739"

/db_xref="taxon:10090"

/clone="2210403N03"

/sex="male"

/tissue_type="stomach"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

142. .1104

/note="unnamed protein product; hypothetical protein (evidence: ProCrest.decoder,NCBI CDS predictor,Longest-ORF)

putative"

/codon_start=1

/protein_id="BAB25910.1"

/db_xref="GI:12843239"

/translation="MTAWDGVLPFPQPRHAASFSVPLLIIVLVLAAASFILPGIRGHSRWFVLRVLLSLFAGAIIVAVHFGDFVFEVWNTSYKAFSPSRVQVHVGLHVLGAVNIFLGTPTQQQLNETIDYNERFTWRLEDYTYKVFHAFKGLPDPVLYLAEKFTSPSPGLVHQYHLPGHYAATLVGVGILLLEDRQCAALHARPTLRLGLFAHRCCLRALRCLFRLLDFQRAALPLPPLRLRLHALLRLLLLTATGILSLGGAVALHVTTPS

CDS

Db 1064 AGTTGAAGAGTCCAGACTTAA--TATTACCACTCTCCTGTGAAATAACTGACTCGG 1120

1238 5'CTGCGTTCCCCCGTCCGACATCGAGGCGGGGAGCAGTGCCCGCCAGCCCTGGGCA 1297

Db
1121 ATTTCTACCCCTCTTTGGGACCCCATAGACTGTGAAGCGCGTGCAGG 1180
1298 CCACACCTCCACCAACGCCCACTGAGAGCTCTGGGCGAGGCGCTGGACATCCGACGGCA 1357

```

Db      1181 AGCATAGCTTAGTCTGGAGAACTGTTCTTCGCTGCGGACGACGACGAGGGGGCACCCCATGTG 1244
Qy      1358 CC-----ACGGAAAGTCTCCTGGGCGGATCTCTGTAATAAAACCTTTTCTTTTGT 1408
Db      1241 CCTTTGCTCTAGAGATGGATTCTTCAGGAGAACTCTAAATGAACCTTTTGTTCCTTT 1300
Qy      1409 TTTT 1413
Db      1301 TTTT 1305

RESULT 3
AI821606/c
LOCUS
DEFINITION
AI821606
537 bp mRNA linear EST 13-DEC-1999
nk08a11.x5 NCI CGAP Co2 Homo sapiens cDNA clone IMAGE:1012892 3'
similar to contains TAR1.t3 TAR1 MER22 repetitive element ;, mRNA
sequence.
AI821606
AI821606.1 GI:5440685
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 537)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other_ESTs: nk08a11.y5
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: National Cancer Institute, Cancer Genome
Anatomy Project (CGAP), Tumor Gene Index
This read has been verified (found to hit its original self in the
correct orientation)
Insert Length: 741 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 459.
Location/Qualifiers
1..537
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1012892"
/tissue_type="tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP Co2"
/note="Organ: colon; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence
5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb.

```

ORIGIN

REFERENCE 1 (bases 1 to 510)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1487 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 476.
Location/Qualifiers
1. 510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1912667"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI-CGAP-Co8"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 35.4%; Score 510; DB 9; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.3e-59;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 904 TTCGGCTCTTCGGGGTCTTCGGCTTGGCTTCCATCTCTAGTGGCGCTGCGCTGCGCGCTC 963
DB 510 TTCGGCTCTTCGGGGTCTTCGGCTTGGCTTCCATCTCTAGTGGCGCTGCGCTGCGCGCTC 451

QY 964 CGCTAGGCTCTCCGGCTCACCCTAGTACGGCGCGCGCTTCTGGTCCAGCTGGCA 1023
DB 450 CGCTAGGCTCTCCGGCTCACCCTAGTACGGCGCGCGCTTCTGGTCCAGCTGGCA 391

QY 1024 ACCGGCTCTGTGGCTCTTCTCGAGGGCGCGCTGGTGGTCTCCAGTATGTTGGCGCC 1083
DB 390 ACCGGCTCTGTGGCTCTTCTCGAGGGCGCGCTGGTGGTCTCCAGTATGTTGGCGCC 331

QY 1084 AGGCTCTTCGACCCCTTCTGGACCAAGCGCCAGGACTGCGCCAGGAGAGAGGGGC 1143
DB 330 AGGCTCTTCGACCCCTTCTGGACCAAGCGCCAGGACTGCGCCAGGAGAGAGGGGC 271

QY 1144 TCACCTCTTATCCTCGGCGACCCCTGCAAGCAGGCGCGCTCTCCAGACTTAAATGT 1203
DB 270 TCACCTCTTATCCTCGGCGACCCCTGCAAGCAGGCGCGCTCTCCAGACTTAAATGT 211

QY 1204 ATCACACTAACCTGTAGGGGGACCAATCTGGACTCTTCCCGCTTGGGACATCGC 1263
DB 210 ATCACACTAACCTGTAGGGGGACCAATCTGGACTCTTCCCGCTTGGGACATCGC 151

QY 1264 AGGCGGGAAGCAGTGCCTCGGCGCTGGCGCAGGAGCTCCAGGAAGGCACTGAGC 1323
DB 150 AGGCGGGAAGCAGTGCCTCGGCGCTGGCGCAGGAGCTCCAGGAAGGCACTGAGC 91

QY 1324 GCTGTGGCGGAGGCTCGGACATCCGAGGCAACAGGGAAAGTCTCTGGGCGATCT 1383
DB 90 GCTGTGGCGGAGGCTCGGACATCCGAGGCAACAGGGAAAGTCTCTGGGCGATCT 31

QY 1384 GTAAATAAACCTTTTTCCTTTTGTGTTTTT 1413
DB 30 GTAAATAAACCTTTTTCCTTTTGTGTTTTT 1

RESULT 6
AI791844 505 bp mRNA linear EST 13-DEC-1999
LOCUS nk08all.y5 NCI-CGAP-Co2 Homo sapiens cDNA clone IMAGE:1012892 5,
DEFINITION similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AI791844
VERSION AI791844.1 GI:5339486
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 505)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: nk08all.s1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: see original entry for original citation
information
This 5' resequenced clone has no previous 5' data to verify this
new read against
Putative full length read
The vector to vector length is 671
Insert Length: 741 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
1. 505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1012892"
/tissue_type="tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI-CGAP-Co2"
/note="Organ: colon; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence:
5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."

FEATURES
source

ORIGIN
Query Match 33.9%; Score 489.2; DB 9; Length 505;
Best Local Similarity 99.2%; Pred. No. 8.3e-57;
Matches 502; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 736 AAGTTCACACCGAGTAGCCCTTGGCGCTGTACCCAGGATACCATCGGGGACACTAC 795
DB 1 AAGTTCACACCGAGTAGCCCTTGGCGCTGTACCCAGGATACCATCGGGGACACTAC 60

QY 796 GCCTCGGCCACGCTATGGGTGGCGTTCTGCTTCTGGTCTCTCTCAACGTGCTCTCTCC 855
DB 61 GCCTCGGCCACGCTATGGGTGGCGTTCTGCTTCTGGTCTCTCTCAACGTGCTCTCTCC 120

QY 856 ACAGCGGCGCGCTCTACGGAGGCGCTGGCACTGTGACCAACCGAGCGCTTCGGCTCTTC 915
|||||
Db 121 ACAGCGGCGCGCTCTACGGAGGCGCTGGCACTGTGACCAACCGAGCGCTTCGGCTCTTC 180
|||||
QY 916 GGGGTCTTCGCGCTTGGCTCCATCTCTAGCGTGCCTCTGCGCGCTCCGCTAGGCTCC 975
|||||
Db 181 GGGGTCTTCGCGCTTGGCTCCATCTCTAGCGTGCCTCTGCGCGCTCCGCTAGGCTCC 240
|||||
QY 976 TCGCGCTCACCACCTCAGTACGGCGCGCTTCTGGGTACGCTGCAACCGGCGCTCTG 1035
|||||
Db 241 TCGCGCTCACCACCTCAGTACGGCGCGCTTCTGGGTACGCTGCAACCGGCGCTCTG 300
|||||
QY 1036 TGCCTCTCTCCTCGAGGGCGCGTGTGAGTCTCCAGTATGTTCCGCCAGCGCTCTTCGC 1095
|||||
Db 301 TGCCTCTCTCCTCGAGGGCGCGTGTGAGTCTCCAGTATGTTCCGCCAGCGCTCTTCGC 360
|||||
QY 1096 ACCCTTCTGGACCAAGCGCCCAAGGACTGCGAGGAGAGAGGGGCTCACCTCTTATC 1155
|||||
Db 361 ACCCTTCTGGACCAAGCGCCCAAGGACTGCGAGGAGAGAGGGGCTCACCTCTTATC 420
|||||
QY 1156 CTCGGCGACCCACTGCAAGCAGCGCGCTCTCCAGACTTAAATGTATCACCCTAAC 1215
|||||
Db 421 CTCGGCGACCCACTGCAAGCAGCGCGCTCTCCAGACTT-AAATGTATCACCCTAAC 479
|||||
QY 1216 CTGTAGGGGGACCCCAATCTGGACTC 1241
|||||
Db 480 CTGTAGGGGGACCCCAATCTGGACTC 505
|||||

RESULT 7
AI660560/c
LOCUS
DEFINITION
IMAGE:2346231 3' similar to contains TARI.t3 TARI repetitive
element ;, mRNA sequence.

QY 1216 CTGTAGGGGGACCCCAATCTGGACTC 1241
|||||
Db 480 CTGTAGGGGGACCCCAATCTGGACTC 505
|||||

QY 1216 CTGTAGGGGGACCCCAATCTGGACTC 1241
|||||
Db 480 CTGTAGGGGGACCCCAATCTGGACTC 505
|||||

QY 1216 CTGTAGGGGGACCCCAATCTGGACTC 1241
|||||
Db 480 CTGTAGGGGGACCCCAATCTGGACTC 505
|||||

QY 1216 CTGTAGGGGGACCCCAATCTGGACTC 1241
|||||
Db 480 CTGTAGGGGGACCCCAATCTGGACTC 505
|||||

QY 1216 CTGTAGGGGGACCCCAATCTGGACTC 1241
|||||
Db 480 CTGTAGGGGGACCCCAATCTGGACTC 505
|||||

dieck@im.wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo.

ORIGIN
Query Match 33.5%; Score 482.4; DB 9; Length 486;
Best Local Similarity 99.4%; Pred. No. 7e-56;
Matches 483; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 943 AGCGTGGCGCTCTGCCGCTCCGCTAGGCTCTCCGCGTCACTCAGTACGGCGCC 1002
|||||
Db 486 AGCGTGGCGCTCTGCCGCTCCGCTAGGCTCTCCGCGTCACTCAGTACGGCGCC 427
|||||
QY 1003 GCCTTCTGGGTACAGCTGGCAACCGCGCTCTCTCTCGAGGGGCGCTGCTG 1062
|||||
Db 426 GCCTTCTGGGTACAGCTGGCAACCGCGCTCTCTCTCGAGGGGCGCTGCTG 367
|||||
QY 1063 AGTCTCCAGTATGTTCCGCCAGCGCTCTTCGACCCCTTCTGACCAAAAGCGCAAGGAC 1122
|||||
Db 366 AGTCTCCAGTATGTTCCGCCAGCGCTCTTCGACCCCTTCTGACCAAAAGCGCAAGGAC 307
|||||
QY 1123 TGAGCCAGGAGAGAGGGGCTCACCTCTTATCTCGGACCCCACTGCACAGGCGCC 1182
|||||
Db 306 TGAGCCAGGAGAGAGGGGCTCACCTCTTATCTCGGACCCCACTGCACAGGCGCC 247
|||||
QY 1183 GCTCTCCAGACTTAAATGTATCACCCTAATCTGAGGGGACCCCAATCTGGACTCC 1242
|||||
Db 246 GCTCTCCAGACTTAAATGTATCACCCTAATCTGAGGGGACCCCAATCTGGACTCC 187
|||||
QY 1243 TTCCCGCGCTTGGGACATCGCAGCGCGGGAAGCAGTGCCTGGCGCGCTGGCCAGGAGA 1302
|||||
Db 186 TTCCCGCGCTTGGGACATCGCAGCGCGGGAAGCAGTGCCTGGCGCGCTGGCCAGGAGA 127
|||||
QY 1303 GCTCAGGAAGGGCACTGAGCGCTCTGCGCGGAGGCGCTCGGACATCCGAGGACCCAGG 1362
|||||
Db 126 GCTCAGGAAGGGCACTGAGCGCTCTGCGCGGAGGCGCTCGGACATCCGAGGACCCAGG 67
|||||
QY 1363 GAAAGTCTCTGGGCGGATCTGTAAATAAACCTTTTCTTTTGTGTTTTTAAAAA 1422
|||||
Db 66 GAAAGTCTCTGGGCGGATCTGTAAATAAACCTTTTCTTTTGTGTTTTTAAAAA 7
|||||
QY 1423 AAAAAA 1428
|||||
Db 6 AAAAAA 1

RESULT 8
BY708714
LOCUS
DEFINITION
musculus cDNA clone 2210403N03 5', mRNA sequence.
ACCESSION
BY708714
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

QY 1423 AAAAAA 1428
|||||
Db 6 AAAAAA 1

QY 1423 AAAAAA 1428
|||||
Db 6 AAAAAA 1

QY 1423 AAAAAA 1428
|||||
Db 6 AAAAAA 1

RESULT 8
BY708714
LOCUS
DEFINITION
musculus cDNA clone 2210403N03 5', mRNA sequence.

ACCESSION
BY708714
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS

QY 1423 AAAAAA 1428
|||||
Db 6 AAAAAA 1

QY 1423 AAAAAA 1428
|||||
Db 6 AAAAAA 1

BY708714
BY708714
musculus cDNA clone 2210403N03 5', mRNA sequence.
BY708714
BY708714.1
GI:27119906
Mus musculus (house mouse)
Mus musculus
Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1014)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 446) NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -40UP from Gibco. Location/Qualifiers 1..446 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:2520674" /tissue_type="colonic mucosa from 3 patients with Crohn's disease" /lab_host="DH10B (phage-resistant)" /clone_lib="Soares Dieckgraebe colon NHCD" /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGCTCTTTTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraebe (Washington University, dieck@im.wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Bonaldo.

ORIGIN

Query Match 31.0%; Score 446; DB 9; Length 446; Best Local Similarity 100.0%; Pred. No. 6.1e-51; Mismatches 0; Indels 0; Gaps 0; Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 971 GTCTCTCCGCGTCACTCACTAGTACGGCGCGCTTCTGGTCACTGGCAACCGGCG 1030
|||||
DB 446 GTCTCTCCGCGTCACTCACTAGTACGGCGCGCTTCTGGTCACTGGCAACCGGCG 387

QY 1031 TCTGTGCTCTTCTCTCGGAGGGCGGTGGTGTCTCCAGTATGTTCCGCCAGCGCTC 1090
|||||
DB 386 TCTGTGCTCTTCTCTCGGAGGGCGGTGGTGTCTCCAGTATGTTCCGCCAGCGCTC 327

QY 1091 TTCCGACCTTCTGGACCAAGCGCAAGGACTGACGCCAGGAGAGAGGGGCTCACCTC 1150
|||||
DB 326 TTCCGACCTTCTGGACCAAGCGCAAGGACTGACGCCAGGAGAGGGGCTCACCTC 267

QY 1151 TTATCTCGGGACCTGCAACAGCAGCGCGCTTCTCCAGACTTAAATGTATCACCA 1210
|||||
DB 266 TTATCTCGGGACCTGCAACAGCAGCGCGCTTCTCCAGACTTAAATGTATCACCA 207

QY 1211 CTAACTGTGAGGGGACCCAAATCTGGACTCTTCTCCCGCTTGGGACATCGCAGGCGG 1270
|||||
DB 206 CTAACTGTGAGGGGACCCAAATCTGGACTCTTCTCCCGCTTGGGACATCGCAGGCGG 147

QY 1271 GAAGAGTGCCCGCCAGGCTGGCCAGGAGAGCTCCAGGAGGSCACTGAGCGTGTCTG 1330
|||||
DB 146 GAAGAGTGCCCGCCAGGCTGGCCAGGAGAGCTCCAGGAGGSCACTGAGCGTGTCTG 87

QY 1331 GCGGAGGCTCTGGGACATCCGACGACCCAGGGAAGTCTCTCGGGCGGATCTGTAAATA 1390
|||||
DB 86 GCGGAGGCTCTGGGACATCCGACGACCCAGGGAAGTCTCTCGGGCGGATCTGTAAATA 27

QY 1391 AACCTTTTCTTTTCTTTTAA 1416
|||||
DB 26 AACCTTTTCTTTTCTTTTAA 1

RESULT 10
AA573825/c
LOCUS
DEFINITION
AA573825
AA573825.1 GI:2348340
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D. CDNA Library Arraying: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Insert length: 741 Std Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 456. Location/Qualifiers 1..466 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:1012892" /tissue_type="tumor" /lab_host="SOLR {kanamycin resistant}" /clone_lib="NCI-CGAP Co2" /note="Organ: colon; Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence: 5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."

AA573825 466 bp mRNA linear EST 12-SEP-1997
nk08all.s1 NCI-CGAP_CO2 Homo sapiens cDNA clone IMAGE:1012892 3', mRNA sequence.
AA573825
AA573825.1 GI:2348340
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D. CDNA Library Arraying: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Insert length: 741 Std Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 456. Location/Qualifiers 1..466 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:1012892" /tissue_type="tumor" /lab_host="SOLR {kanamycin resistant}" /clone_lib="NCI-CGAP Co2" /note="Organ: colon; Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Bulk colon villous adenoma. 5' adaptor sequence: 5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."

ORIGIN

Query Match 30.9%; Score 445.4; DB 9; Length 466; Best Local Similarity 98.5%; Pred. No. 7.1e-51; Mismatches 6; Indels 1; Gaps 1; Matches 460; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 947 TGCCTCTGCGCTCGCGCTAGGCTCTCGCGCTCACCACTCAGTACGGCGCGCCT 1006
|||||
DB 466 TGCCTCTGCGCTCGCGCTAGGCTCTCGCGCTCACCACTCAGTACGGCGCGCCT 407

QY 1007 TCTGGGTCACTGGCAACCGCGTCTCTGCTCTTCTCGGAGGGCGGTGGTGATC 1066
|||||
DB 406 TCTGGGTCACTGGCAACCGCGTCTCTGCTCTTCTCGGAGGGCGGTGGTGATC 347

QY 1067 TCCAGTATGTTCCGCCAGCGCTCTTCGACCCCTTCTGACCAAAGCGCAAGGACTGCA 1126
|||||
DB 346 TCCAGTATGTTCCGCCAGCGCTCTTCGACCCCTTCTGACCAAAGCGCAAGGACTGCA 287

QY 1127 GCCAGGAGAGGGGGTCACTCTTATCTCTGGCGACCTGCAACAGAGGGCGCTC 1186
|||||
DB 286 GCCAGGAGAGGGGGTCACTCTTATCTCTGGCGACCTGCAACAGAGGGCGCTC 227

QY 1187 TCCAGAGCTTAAATGTATCACCACTAACCTGTGAGGGGAGCCCAATCTGACTCTCTCC 1246
|||||
DB 226 TCCAGAGCTTAAATGTATCACCACTAACCTGTGAGGGGAGCCCAATCTGACTCTCTCC 167

QY 1247 CCGCCCTGGGACATCGCAGGCGCGGAGAGAGTGCCTCCGCGAGGCTGGCCAGGAGCTC 1306
 Db 166 CCGCCCTGGGACATCGCAGGCGCGGAGAGAGTGCCTCCGCGAGGCTGGCCAGGAGCTC 107
 QY 1307 CAGGAGGCGCACTGAGCGCTGCTGGCGGAGGCGCTCGGACATCCGCGAGGCGAGGAGAA 1366
 Db 106 CAGGAGGCGCACTGAGCGCTG-TGGCGGAGGCGCTCGGACATCCGCGAGGCGAGGAGAA 48
 QY 1367 GTCTCTGGGCGATCTGTAATAAATTAACCTTTTCTTTTCTTTTCTTTT 1413
 Db 47 GTCTCTGGGCGATCTGTAATAAATTAACCTTTTCTTTTCTTTTCTTTT 1
 RESULT 11
 AI991272/c
 LOCUS
 DEFINITION wu41h04.x1 Soares_Dieckgraeft_colon_NHCD Homo sapiens cDNA clone
 IMAGE:2522647 3', mRNA sequence.
 ACCESSION AI991272
 VERSION AI991272.1 GI:5838177
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 434)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 558 Std Error: 0.00
 Seq primer: -40UP from Gibco.
 Location/Qualifiers
 1. 434
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2522647"
 /tissue_type="colonic mucosa from 3 patients with Crohn's
 disease"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Soares Dieckgraeft_colon_NHCD"
 /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization. Tissue samples
 provided by Dr. Brian Dieckgraeft (Washington University,
 dieck@im.wustl.edu); colonic mucosa represents a range of
 disease involvement from moderate to severe Crohn's
 disease; samples include both perforating (fistulas) and
 non-perforating samples. Library constructed by Bento
 Soares and M. Fatima Bonaldo."
 ORIGIN
 Query Match 30.1%; Score 434; DB 9; Length 434;
 Best Local Similarity 100.0%; Pred. No. 2.5e-49;
 Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 981 GCTCACCACCTCAGTACGCGCGCGCTTCTGGGTACGCTGGCAACCGCGTCTGTGCT 1040
 Db 434 GCTCACCACCTCAGTACGCGCGCGCTTCTGGGTACGCTGGCAACCGCGTCTGTGCT 375
 QY 1041 CTTCCTCGGAGGCGCGTGGTGTAGTCTCCAGTATGTTCCGCCAGCGCTCTTCGCACCCCT 1100
 Db 374 CTTCCTCGGAGGCGCGTGGTGTAGTCTCCAGTATGTTCCGCCAGCGCTCTTCGCACCCCT 315

QY 1101 TCTGGACCAAGAGCGCCCAAGGACTGCAGCCAGGAGAGAGGGGCTCACCTCTTATCTCTCGG 1160
 Db 314 TCTGGACCAAGAGCGCCCAAGGACTGCAGCCAGGAGAGAGGGGCTCACCTCTTATCTCTCGG 255
 QY 1161 CGACCCACTGCACAAGCAGGCGCGCTCTCCCGAGCTTAAATATGATATCACCACCTAACCTGTG 1220
 Db 254 CGACCCACTGCACAAGCAGGCGCGCTCTCCCGAGCTTAAATATGATATCACCACCTAACCTGTG 195
 QY 1221 AGGGGGACCCCAATCTGGACTCTCTCCCGCTTCCCGAGCATCGCAGGCGCGGAGCAGTGC 1280
 Db 194 AGGGGGACCCCAATCTGGACTCTCTCCCGCTTCCCGAGCATCGCAGGCGCGGAGCAGTGC 135
 QY 1281 CCGCCAGGCGCTGGGCGCAGGAGAGCTCCAGGAAGGGCACTGAGCGCTGTGGCGCAGGCGC 1340
 Db 134 CCGCCAGGCGCTGGGCGCAGGAGAGCTCCAGGAAGGGCACTGAGCGCTGTGGCGCAGGCGC 75
 QY 1341 TCGGACATCCGAGGCGCAGGAGAAAGTCTCTCTGGGCGGATCTGTAAATAAACCTTTT 1400
 Db 74 TCGGACATCCGAGGCGCAGGAGAAAGTCTCTCTGGGCGGATCTGTAAATAAACCTTTT 15
 QY 1401 TCTTTTGTTTT 1414
 Db 14 TCTTTTGTTTT 1
 RESULT 12
 AA593860/c
 LOCUS
 DEFINITION nn19f03.s1 NCI_CGAP_Col2 Homo sapiens cDNA clone IMAGE:1084349 3',
 mRNA sequence.
 ACCESSION AA593860
 VERSION AA593860.1 GI:2408538
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 414)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Stratagene, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1204 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 406.
 Location/Qualifiers
 1. 414
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1084349"
 /sex="mixed"
 /tissue_type="colon tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /clone_lib="NCI_CGAP_Col2"
 /note="Organ: colon; Vector: Bluescript SK-; Site 1:
 EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. Pooled colon tumors. 5' adaptor sequence: 5'
 GAATTCGGCAGCAG 3' 3' adaptor sequence: 5'
 CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.2 kb."
 ORIGIN

	Query Match	28.7%	Score 414;	DB 9;	Length 414;
	Best Local Similarity	100.0%;	Pred. No. 2.3e-46;		
	Matches 414;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	999	CGCCGCCCTCTGGGTACAGCTGGCAACCGGCGTCTCTGTGCTCTCTTCTCGAGGGGCCGT	1058		
Db	414				
QY	1059	GGTGAGTCTCCAGTATGTTGGGCCCAGCGCTCTTTCGCACCCCTTCTGGACCAAGCGCCAA	1118		
Db	354				
QY	1119	GGAATGCAAGCAGGAGAGAGGGGGCTCACCTCTTTATCTCTGGCGACCCACTGCACAAGCA	1178		
Db	294				
QY	1179	GGCCGCTCTCCAGACTTAAATGTATCACCACTAACTGTGAGGGGACCCCAATCTGGA	1238		
Db	234				
QY	1239	CTCCTTCCCGCCTTGGGACATCGCAGGCCGGAAGCAGTGCCTCCCGCCAGGCTGGGCCAG	1298		
Db	174				
QY	1299	GAGAGCTCCAGGAAGGGCACTGAGCGCTGTGGCGCGAGGCTCGGACATCCGCAAGGCAC	1358		
Db	114				
QY	1359	CAGGGAAAGTCTCTGGGGCGATCTGTAAATAAAACCTTTTCTTTTCTTTTCTTTTCTTTT	1412		
Db	54	CAGGGAAAGTCTCTGGGGCGATCTGTAAATAAAACCTTTTCTTTTCTTTTCTTTTCTTTT	1		

RESULT 13	AI983793/c	AI983793	411 bp	mRNA	linear	EST 27-OCT-1999
LOCUS						
DEFINITION		wu20c09.x1 Soares_Dieckgraefe_colon_NHCD				Homo sapiens cDNA clone
		IMAGE:2520592 3'				mRNA sequence.

```

SOURCE      Homo sapiens {human}
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 411)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgaps-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40Up from Gibco
            High quality sequence stop: 400.
FEATURES    location/Qualifiers
            1..411
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:2520592"
            /tissue_type="colonic mucosa from 3 patients with Crohn's
            disease"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="Soares-Dieckgraefe colon NHCD"
            /note="Organ: colon; Vector: pT73b-Pac (Pharmacia) with a
            modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5',
            TGTTACCAATCTGACTGGAGCGCGCGCTTTTTTTTTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT73 vector. Library

```

RESULT 14
AI274929/c

DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE	AUTHORS	TITLE
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

**JOURNAL
COMMENT**

US 0964180213P1



Creation date: 07-08-2004
Indexing Officer: NGOLSON - NA-WAL GOLSON
Team: OIPE Scanning
Dossier: 09641802

Legal Date: 07-07-2004

No.	Doccode	Number of pages
1	SRFW	1
2	SRNT	266

Total number of pages: 267

Remarks:

Order of re-scan issued on